UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author VESPE, Michele

Multi-Perspective Radar Target Classification

Michele Vespe

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of the

University of London.

Department of Electronic & Electrical Engineering

University College London

May 23, 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Michele Vespe, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The problem of radar target classification is examined for the case when more than one perspective or viewing angle of the target is available to the sensor. It is shown how multiple perspectives enhance the classification performance through the analysis of the classification results on full scale target signature measurements. Furthermore, the classifier capabilities are explored as a function of the number of perspectives, the angular perspective displacements, the signal to noise ratio, the resolution and the illuminating wavelength. In order to remove any possible bias that could be introduced by a single individual classifier, various approaches to Multi-Perspective (M-P) classification have been implemented, using both High Range Resolution (HRR) profiles and 2-D images from real Synthetic Aperture Radar (SAR) and Inverse SAR (ISAR) data. The classification performance is also described for those applications where the perspective displacement information is known and can be processed. The results show a consistent improvement in classification performance as the number of perspectives increases. The techniques employed also provide considerable insight into the classification process highlighting the degree of complexity of this extremely challenging problem.

Acknowledgements

I would like to gratefully acknowledge my supervisors Prof. Chris J. Baker and Prof. Hugh D. Griffiths for their insightful suggestions and enthusiastic support. I will forever be indebted to them for the opportunity they gave me.

The work reported in this thesis was funded by the Electro-Magnetic Remote Sensing (EMRS) Defence Technology Centre, established by the UK Ministry of Defence and run by a consortium of SELEX Sensors and Airborne Systems, Thales Defence, Roke Manor Research and Filtronic. I would like to thank Thales Defence for providing ADAS files to investigate HRR profiles from real targets and the punctual and invaluable feedback offered by the EMRS DTC supervisory board during meetings and technical conferences.

I would like to thank the numerous colleagues and friends at the Department of Electronic & Electrical Engineering at University College London for the helpful discussions and inspiring comments.

I would like to thank also my friends Alan and Rhea who taught me aspects of the English culture I would had never learnt by myself. Thanks are due to London and all the people I met in this incredible city.

A special 'thank-you' to my sisters and Paola for having supported and encouraged me notwithstanding the distance. Finally, this work is completely dedicated to my mother, and all she has done for me.

Publications

Parts of the research work presented here have been published:

- M. Vespe, C. J. Baker, and H. D. Griffiths. Multi-Perspective radar target classification. In *EMRS DTC* 1^{st} *Annual Technical Conference*, Edinburgh, May 2004.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Multi-Perspective target classification. In *2005 Int. IEEE Radar Conference*, pages 877-882, Arlington (VA), May 2005.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Radar target classification using multiple perspectives. In *EMRS DTC* 2^{nd} *Annual Technical Conference, Edinburgh*, June 2005.

- M. Vespe, C. J. Baker, H. D. Griffiths, P. Lombardo, and D. Pastina. Feature extraction for SAR target classification. In *London Communications Symposium*, London, 2005.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Node location for netted radar target classification. In *2006 International Waveform Diversity and Design Conference*, pages 256-260, Hawaii, January 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Aspect dependent drivers for Multi-Perspective target classification. In *2006 IEEE Radar Conference*, pages 256-260, Verona (NY), April 2006.

- M. Vespe, C. J. Baker, H. D. Griffiths, P. Lombardo, and D. Pastina. Impact of angular processing techniques and cross-range resolution on 2-D radar target classification. In 6^th *European Conference on Synthetic Aperture Radar*, Dresden, May 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Perspective aperture dependency of netted radar target classification. In *EMRS DTC* 3^{rd} *Annual Technical Conference*, Edinburgh, July 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Multi perspective target classification. In *SET-095 NATO Specialist Meeting on Bistatic and Multistatic Radar and Sonar Systems*, La Spezia, July 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Frequency diversity vs large bandwidth reconstruction: information content for netted sensor ATR using ISAR images. In *International Conference on SAS and SAR*, Lerici (Italy), September 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Outline structural representation for radar target classification based on non-radar templates. In *CIE Int. Radar Conference*, October 2006.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Radar ATR using Multi-Frequency Illuminations. *Special Issue of IET: Radar, Sonar and Navigation*, in revision.

- M. Vespe, C. J. Baker, and H. D. Griffiths. Radar target classification using multiple perspectives. *IEE journal paper*, in revision.

- M. Vespe, G. Jones, and C. J. Baker. Echolocating Bats for Autonomous Navigation. *Journal of the Acoustical Society of America*, in revision.

- C. J. Baker, M. Vespe and G. Jones. Diversity strategies: Lessons from natural systems. *Paper submitted to the 2007 International Waveform Diversity and Design Conference.*

Contents

List of Figures

List of Tables

Acronyms

2-P Two-Perspective

3-P Three-Perspective

ADAS Airborne Data Acquisition System

ANN Artificial Neural Network

ATR Automatic Target Recognition

AUC Area Under the Curve

CEM Computational Electro-Magnetics

CCR Correct Classification Rate

DFT Discrete Fourier Transform

EO-IR Electro-Optics Infrared

ESM Electronic Support Measures

FANN Feed-forward Artificial Neural Network

FT Fourier Transform

FFT Fast Fourier Transform

GTD Geometrical Theory of Diffraction

HMM Hidden Markov Model

HRR High Range Resolution

HRRP High Resolution Range Profile

IFF Identification Friend or Foe

IFT Inverse Fourier Transform

IFFT Inverse Fast Fourier Transform

ISAR Inverse Synthetic Aperture Radar

JEM Jet Engine Modulation

K-NN K-Nearest Neighbour

LFM Linear Frequency Modulation

LOS Line Of Sight

MIDAS Mobile Instrumented Data Acquisition System

MLP Multi-Layers Perceptron

M-F Multi-Frequency

M-P Multi-Perspective

MSTAR Moving and Stationary Target Acquisition and Recognition

NCTR Non-Cooperative Target Recognition

PCA Principal Components Analysis

PCR Pulse Compression Ratio

PR Pattern Recognition

PRF Pulse Repetition Frequency

PRI Pulse Repetition Interval

PSF Point Spread Function

RBF Radial Basis Function

RBFNN Radial Basis Function Neural Network

RCS Radar Cross-Section

ROC Receiver Operator Characteristic

ROI Region Of Interest

RF Radio Frequency

SAR Synthetic Aperture Radar

SC Scattering Centre

SCM Scattering Centres Model

S-F Stepped-Frequency

SNR Signal to Noise Ratio

SVA Spatially Variant Apodization

TR Tomographic Reconstruction

ZDC Zero Doppler Clutter

Introduction

Motivation

The ability to reliably detect and classify targets using radar systems has many important applications for both military and civil domains. The military domain is related to Identification of Friend or Foe (IFF) systems and Non-Cooperative Target Recognition (NCTR) functions such as in airborne (fighters and missiles), ground-based or maritime radars for engaging flying, ground or naval targets. The civil contribution of Automatic Target Recognition (ATR) is primarily to reinforce both air and coastal surveillance, nowadays crucial for homeland defence. Air Traffic Control (ATC) could also been supported by automatic recognition tools, reducing the risk of tragic incidents.

Investigations into radar target classification have been fuelled over many years by the progressive increase in the amount and quality of information available from advanced radar systems. Chief amongst these is the development of very high range resolution that provides a detailed signature of the target to be classified. However, as the electromagnetic backscattering behaviour of complex targets is extremely sensitive to radar parameters, orientation and environment, traditional classification techniques based on a single view of the target only provide a limited level of performance. Indeed, this level of performance in real situations is insufficient to enable automation. The main idea of the work of this thesis has been inspired by human and animal behaviours, especially from echolocating mammals (University of Maryland Batlab Website and Simmons et al. (1992)): the use of multiple views to explore the target significantly improves the solution of the identification ambiguities arising from unfavourable aspects.

A network of multistatic radar systems offers a way of collecting backscattered radiation with higher information content from the target than is possible using a single radar. An additional advantage of using multistatic or bistatic systems is to provide a valuable counter to stealth technology which has been largely aimed at defeating monostatic radars only.

This study on potential improvements in target classification using Multi-Perspective views of targets has been formulated to develop ATR for a network of radars. Although the concept of netted radar is not new it is only recently that the development in low cost COTS technologies and high speed digital signal processing that has enabled these advanced and complex concepts to be realised. Additionally, the idea of multi-aspect ATR can be incorporated into a single existing radar system collecting different perspectives from the object as it moves during the time on target.

Sonar imagery may exhibit similar characteristics since the target dimensions and wavelengths are comparable, although to some extent scattering mechanisms are somewhat different.

Aims

In this research work, the impact of a Multi-Perspective processing is examined in terms of radar target classification performance improvements over the traditional Single-Perspective case.

The principal aims of this work can be summarised as follows:

- To evaluate techniques for target image reconstruction from partial information. This can be thought of as the combination of multiple signatures from the target collected either by a network of radars or by a single manoeuvering system or by a static system and a moving object.
- To examine the relationship between the radar target reflectivity function, high resolution and illuminating wavelength on classification performance.
- To develop tools to process Multi-Perspective (M-P) target data.
- To understand classification improvements made possible through the use of multiple

signatures of the target, collected using different aspect angles and/or illuminating frequencies.

- To investigate the information content in terms of classification contained in multiple perspectives and their angular separation.

Brief History

In this section we are not attempting to give a complete and detailed overview of the evolution of radar systems (exhaustively treated by Swords (1986)) but only a list of significative milestones in its history.

Radar (*RAdio Detection And Ranging*) is a system that is able to determine the reflectivity of targets by transmitting electromagnetic waves and subsequently sensing the backscattered returns (Page, 1962). Heinrich Hertz in 1886 described for the first time how to radiate radio waves by varying the electric current in a wire. He also found that some materials reflected radio waves. At the end of the nineteenth century, Lodge (England), Branly (France), Popov (Russia) and Thomson (USA) detected radiation from electrical discharges. In particular, Popov was the first who used a single antenna both in transmission and reception of radio waves.

It was in 1904 that Christian Hülsmeyer demonstrated his radar-like system (Telemobiloskop) on a ships in the harbour of Rotterdam.

In 1922, Guglielmo Marconi made a speech describing how radio waves could be used for detecting objects. Meanwhile, the detection of wooden ships using Continuous Waveform (CW) was experimentally demonstrated by the American scientists Albert H. Taylor and Leo C. Young at the Naval Research Laboratory (NRL). The advent of pulse radar was encouraged by the need for range resolution to locate the detected target, impracticable using the original CW radars. In 1935, using a pulse radar, Sir Robert Alexander Watson-Watt developed the first aircraft locating radar, able to detect bomber aircrafts at ranges greater than 60 km.

In early 1940 Sir John Randall and Henry A. H. Boot of Birmingham University developed the cavity magnetron, a small device which generated microwave frequencies more efficiently than previous devices, enabling high transmission powers. Practical centimetric radars

were therefore developed, allowing for the detection of smaller objects and the use of smaller antennas which could be carried by aircrafts. During the World War II, the first radar systems were deployed. Great Britain, Germany, France and the United States all used radar, which was also called "magic eye", a device that could surmount the darkness, fog and weather to navigate their ships, airplanes, and detect the enemy before being attacked. The "chain home" system, a ground based network of radars, is often credited with allowing the British Royal Air Force (RAF) to defeat the much larger German Luftwaffe. Since World War II, with the use of Doppler information, the development of radar technologies for both military and civil applications has led to unlimited use of this device for safety, detection, navigation, tracking and weather analysis.

Improvements in resolution have led to increasing information content of the backscattered signatures. This has spawned interest in classifying radar targets which has become a progressively more important application together with the already well-established detection and localisation techniques. This is a consequence of the improved ability to resolve the target into multiple resolution cells, providing a more detailed backscattering description.

Literature Review

In this section, the key concepts of the thesis are introduced. They are further discussed and integrated in the following chapters.

The ability to detect and locate targets on a day/night, all weather basis, over wide areas, has long made radar a key sensor in many military and civilian applications. It is well recognised that the utility of the information supplied by a radar system would be hugely enhanced if targets could additionally be classified. This might be by type (e.g. to differentiate helicopters from civilian airliners from military jets) or even better, within type (e.g. a Chieftain tank from a T32 tank). In addition, huge quantities of imagery may be produced by SAR systems potentially placing unrealistic demands on a human image interpreter. This explains why it would be even more advantageous if classification could be automated, or at least partially automated. Algorithmic aids are necessary if the data are ever to be fully exploited (or even reviewed). However, except for a few niche applications, it has proved far from straightforward

to create robust, reliable classifiers able to cope with a wide range of target and clutter conditions. When a radar system illuminates a target, the energy is reflected in all directions with some of this signal returning back in the direction of the radar. The reflected signal may be modified due to a combination of the target itself, the clutter in the illumination cell and often a forward scattering component due to multipath. The classification system attempts to recognise the target based upon an analysis of this reflected signal. It is clear that the contributions to the reflected signal from anything other than the target immediately begin to considerably complicate the problem. This becomes even more severe as, due to the coherency of radar systems, there are fluctuations caused by any slight change in geometry due to scatterers moving in and out of phase. Additionally, of course, there is a huge range of possible targets that have to be distinguished from one another. Chiefly for these reasons, reliable target classification has remained somewhat elusive.

The main contributions and the progress achieved in ATR are now briefly listed, distinguishing the three different disciplines involved in the classification process: the nature of data to be processed, the pattern recognition approach used and the direction followed in order to enhance the trustworthiness and robustness of the performance.

Data Types. Amongst many other techniques discussed in Chapter 2, where a more detailed literature survey is discussed on this topic, radar ATR can be based either on 1-D, 2-D, 3-D or interferometric signatures. The point-scatterer assumption, the concept of range resolution and the Scattering Centre Model (SCM) developed by Rihaczek and Herschowitz are discussed in Chapter 1. ATR based on 1-D imagery of the target is often employed due to its simplicity in terms of implementation and signal processing and the fact that the profile is less vulnerable to motion induced corruption. However, the use of 1-D profiles in an automatic classifier may lead to performance that is prone to low classification rates due to the weak robustness to noise (Kim et al., 2001), aspect angle (Wu et al., 2002) and additional inclusion of a clutter component that is scenario dependent (Williams et al., 2000). 2-D imagery is less likely to contain clutter (although there will still be a significant multipath or forward scattering component) and hence would seem

to promise a more detailed and authentic (but nevertheless still corrupted) representation of the target backscattered signature (Novak, 1991). It is also still prone to unavoidable sensitivity to illumination angle. 2-D imagery also requires more sophisticated signal processing and if the data are collected over an extended period of time (as in a Stepped-Frequency radar system) correction for irregular target movement is required if de-focusing is to be avoided. Such correction can be far from straightforward, especially as the detailed form of the motion corruption is usually unknown (Miller et al., 2004). Alternatively, when the target has fast moving parts as the engine blades of an aircraft or the blades of helicopters, the time series of frequency domain signatures can be modelled (e. g. Tait (2006), Martin and Mulgrew (1990)) and the Jet Engine Modulation (JEM – originated by the work by Prof. William Bardo at RSRE Malvern) measurements can be used as a pattern to be classified (Bell and Grubbs (1993), Piazza (1999)). Helicopter recognition (e.g. Misiurewicz et al. (1997)) is the less aspect-independent ATR approach, whereas for any other radar object recognition approach (i.e. using 1-D, 2-D and JEM measurements) the classification performance is highly subject to the particular direction sensing the target.

Classification Approaches. After the target is detected and its signature pre-processed, ATR can be approached using three fundamental methodologies:

- *Template-Based*: the database consists of previously collected measurements from targets of interest to form the template library (a series of class labelled patterns) while the recognition consists of finding the closest class producing similar returns (e.g. Ikeuchi et al. (1996)). Although rather simple to implement, this approach is significantly dependent on the radar system and specification, as well as on the particular conformation of the target representatives.
- *Computational Electro-Magnetics* (CEM): the target is is designed using CAD models and the radar signature subsequently predicted (e.g. FACETS^{TM}, X-Patch®, FEKO©). The possibility of modelling any target class having any possible configuration are the principal advantages of this method, counterbal-

anced by the difficulty of obtaining reliable simulated signatures.

- *Model-Based*: through the *a priori* knowledge of the target physical and electrical properties it is possible to detect those features that discriminate different classes in the feature space (e.g Moses and Potter (1997), Chiang et al. (2000)). The advantage of model-based approaches resides in potentially obtaining a library more robust than using CEM maintaining the benefit of not necessarily requiring radar measurements to represent the target class.
- *Multi-Approach*: a mixture of the methodologies described above can be done, using different sources of information and a number of different models that can be combined using Bayesian Networks (Tait, 2006).

The approaches for the template library construction and its impact on classification are discussed in detail in Chapter 4.

The Pattern Recognition algorithm to be used in target classification is a well-known topic of discussion in the radar ATR community. The primary aim is to maximise the probability of correct classification (Mitchell and Westerkamp, 1999), although a more detailed list of classification requirements is discussed in Chapter 4. *Correlation techniques* have been implemented in order to establish the similarity between templates and input vectors (e.g. Hudson and Psaltis (1993)). *Statistical* methods based on the Bayes' Theorem (e.g. De Vore and O'Sullivan (2002); Zyweck and Bogner (1996)), *non-parametric* (K-Nearest Neighbour, e.g. Miller et al. (2004); Schumacher and Schiller (2005)), *associative mapping* (multilayer neural networks, e.g. Gross (1999)) and *non-metric* techniques (e.g syntactic approaches as in Sands and Garber (1989) and in Vespe et al. (2006a)) have also been investigated for radar ATR.

It is also worth noting that, although analysed by Cohen (1991), there are no rigourous and universally accepted definitions for the terms "identification", "classification", and "recognition". Their meaning also slightly changes according to different communities (i.e. operational or R&D).

A possible definition, suggested by the authors, is described here (figure 1):

- *Classification*: after detection and discrimination (concepts described in Chapter 2), the sensed attributes from the target are processed in order to make a first decision on the type of object (e.g. helicopter, aircraft) and its domain (military or civil) and tasks (fighter, reconnaissance, etc...).
- *Recognition*: after class labelling the target, its particular model can be recognised as well as its particular version (e.g. loads and weapons).
- *Identification* : the ultimate stage of the process is the serial number identification, which characterises a unique exemplar. This phase is usually aided by cooperative techniques.

In this work the terms "classification" and "recognition" are used interchangeably to address the ability of the radar system and subsequent processing to determine a category to which a target belongs.

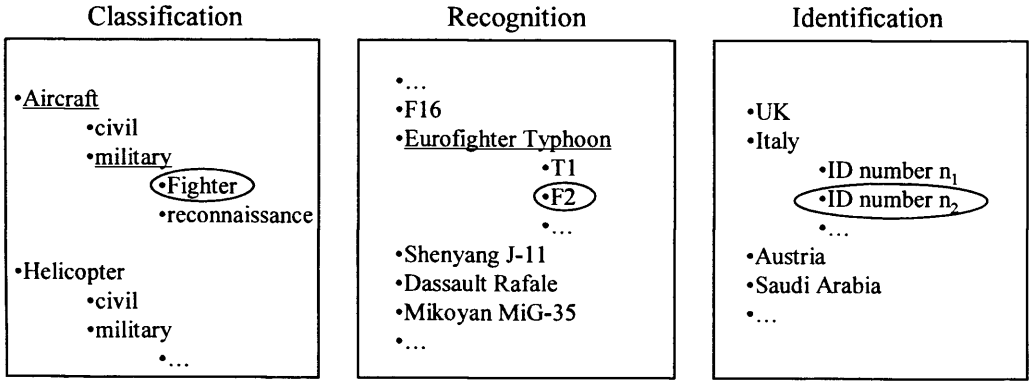

Figure 1: *"Classification", "Recognition" and "Identification".*

Information Content for ATR and Multiple Perspectives. To perform ATR in a useful manner it is necessary to achieve a sufficiently high degree of confidence that a target has been correctly classified and hence it is extremely important to minimise the probability of misclassification. This is because in both military and civil applications there are likely to be safety of life aspects that demand an acceptably high level of performance whose limits and variability are fully understood.

Most of the approaches to classification have concentrated on exploiting higher and higher spatial resolutions both in one (e.g Hu and Zhu (1997)) and two (e.g. Novak et al. (1997)) dimensions, or by synthetically increasing the resolution capabilities by means of superresolution postprocessing (Novak et al. (1996), Tsao and Steinberg (1988), Schmidt (1979), Bienvenu and Kopp (1980)) and sidelobe reduction techniques (Stankwitz and Kosek (1995), Stankwitz et al. (1996)). This is because the actual location of the main scattering centres is crucial for reliable classification. The more detailed their spatial density distribution description, the more accurate and unambiguous the information content to be exploited. Research has also been reported examining the role of polarimetric information (e.g. Sadjadi (2002), Novak et al. (1997), Novak and Hesse (1992)) often in conjunction with high resolution, demonstrating that the polarimetric information provides for significant discrimination between a number of structures, with the result of aiding the data classification.

However, there has been little research published that examines the utility of angular diversity for improving classification performance (Shihao et al., 2005; Xuejun et al., 2002; Runkle et al., 1999a; Salazar et al., 2002). This is a little surprising as it is clear (i) there must be additional information in multiple perspectives and (ii) it is extremely easy for existing systems to acquire such data. A simple example would be an aircraft flying past a scene of interest, which will automatically generate multiple viewing angles. In fact, to date, the variation of target signature with angle has been a source of performance degradation in single perspective classification approaches. Alternatively, networks of co-operating radar systems (Baker and Hume, 2003) offers a possible solution for improving classification performance. These have the added advantages of improved system sensitivity, tailored coverage and reduced vulnerability to electronic and physical attack.

The early stage of Multi-Perspective classification was based on the radar-target perspective change within the time on target and it was related to successive measurements of the backscattering as reported by Jacobs and O'Sullivan (2000), Liao et al. (2004),

Pei and Bao (2005), Salazar et al. (2002) and Liao et al. (2001). In the last three works, the authors applied Hidden Markov Models (HMM), well known in speech recognition (Rabiner, 1989), to multi-aspect classification. 1-D radar and sonar signatures were used as well as 2-D images subsequently subdivided into a sequence of sub-images. The principal issue with their approach was that the advantages brought by multiple looks at the target deteriorate with the correlation of the signatures (as it will be explained in Chapter 5 and 6), which is large for relatively small perspective displacements between consecutive looks.

The HMM approach, used to model different states (target orientations), was then extended to non-contiguous perspectives, modelling the HMM transitions in relation to the angular perspective displacement between states as discussed by Shihao et al. (2005). They also propose an algorithm to select adaptively the optimal perspective separation between looks based on minimising the predicted entropy of the decision, and show the results on a set of sonar targets. Their contribution, carried out contemporaneously with the research reported here, provides a further insight into the classification performance improvement through the use of multiple perspectives, although confined to a set of sonar targets showing a relatively simple shape conformation and a number of symmetries. This aspect has a significant impact on multi-aspect recognition. The HMM multi-aspect approach has been implemented and compared with the one developed by the authors. The performance analysis is shown in Chapter 5 and highlights that, although the angular perspective displacement improves the M-P classification capabilities, the state-representation needs to be more efficiently designed. Nevertheless, the conclusions brought by Shihao et al. can be integrated with this work, being more focused on the aspect dependent scattering behaviour of complex targets rather than the particular pattern recognition approach used. The angular perspective dependency of the network topology, as well as the Signal-to-Noise-Ratio (SNR) effects on Multi-Perspective (M-P) ATR are described. In addition, different classification methodologies and feature extraction algorithms applied in order to evaluate the extent of informa-

tion content enhancements given by multiple views of the target. The objects of interest (real measurements from complex targets) are also investigated in more detail since the analysis of M-P classification cannot leave this out of consideration: the results of M-P classification discussed for different dataset will show the degree to which the perspective displacement is highly dependent on the symmetries cross-affinities between the classes of the population considered.

Finally we focus on the application of the concepts of classification "*reliability*", "*unknown*" class and "*test variant*" input targets, necessary aspects to evaluate the classification performance in a realistic way and diffusely discussed later on.

The number of approaches followed to perform radar target classification, the different techniques employed to extract backscattering information from real data, and the instruments used to analyse the classification performance all corroborate the main finding of this research work, represented by the description of the Multi-Perspective classification benefits over the single perspective case.

Layout of the Thesis

The first chapter is related to high resolution techniques to produce one-dimensional and two-dimensional signatures from radar targets, as well as their possible reconstruction using multiple perspectives (i.e. multiple signatures obtained by sensing the target from different perspectives). This is followed by an overview on ATR and feature extraction techniques in Chapter 2. The data processing procedures to form 1-D and 2-D signatures are then implemented using real measurements (Chapter 3), and the resolution concept further described. The analysis of the real datasets parameters and the traditional single perspective results are described in Chapter 4. Multi-Perspective classification is first attempted in a "disjoint" mode, i.e. neglecting the angular displacements between the aspects sensing the object (Chapter 5) and subsequently analysed using a "joint" M-P classifier after estimating the perspective topology of the problem (Chapter 6). In Chapter 7, an insight into the information content of radar target backscattering is provided, focusing on the low level point-like features due to surface discontinuities, in contraposition to the stronger returns from Scattering Centres (SCs) tradi-

tionally described as the main source of backscattering information. The point-like returns of the illuminated near regions of the target are investigated for optical-based radar ATR. The effects of high resolution and frequency diversification are analysed in Chapter 8 as a further backscattering information investigate and evaluate. Finally, the findings of this research are summarised and conclusions drawn.

Chapter 1

High Range Resolution and Radar Imaging

This chapter reviews the main properties of range profiles as well as the concept of "resolution" and the processing techniques used to form one- and two- dimensional signatures. The results of the implementation of such processing are eventually shown on real data in Chapter 3.

Assuming a single point scatterer (figure 1.1), the pulse transmitted by a monostatic radar system is reflected back and subsequently received after a time delay τ during which the pulse covers the two-way path $2r$ between the radar and the scatterer.

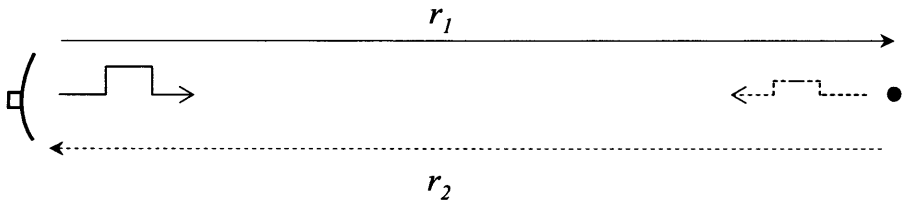

Figure 1.1: *Two-way propagation delay. The transmitted pulse (solid line) covers the path* $r = r_1 = r_2$*, is then scattered back (dashed line) and received after a time* τ*. The received power is expressed by the radar equation.*

In case of monostatic systems, the range $r = r_1 = r_2$ can be therefore expressed as follows:

$$r = \frac{c\tau}{2} \tag{1.1}$$

where c is the speed of light. The minimum distance Δr that could be resolved in range (*range resolution*), is therefore related to the pulse length T_p and is exhaustively discussed in

the following sections.

The received power P_{rx} is related to the radar equation (Skolnik, 1990), which gives the relationship between the received signal strength and the radar characteristics:

$$P_{rx} = \frac{P_{tx}}{4\pi r_1^2} \cdot G \cdot \frac{\sigma}{4\pi r_2^2} \cdot A_{e,rx} \tag{1.2}$$

Equation 1.2 tells that the received signal strength of a target having a Radar Cross Section (RCS, see Section 3.1) σ is proportional to the transmitter antenna gain G and the receiving antenna's effective area $A_{e,rx}$. Besides, the received power is reduced by the two-way spreading loss $4\pi r^2$ due to the spreading of the power over the spheres of radius r_1 and r_2 surrounding the transmitting and receiving antenna. In case of monostatic systems, the receiver and the transmitter are co-located, therefore $r_1 = r_2$, and the gain $G = 4\pi A_e/\lambda^2$ is equal both for the antennas:

$$P_{rx} = \frac{P_{tx}\, G^2\, \lambda^2\, \sigma}{r^4\, (4\pi)^3} \tag{1.3}$$

As illustrated in figure 1.2, a range profile is a representation of the illuminated 3-D spatial scatterer distribution mapped accordingly to their distance to the sensing system. Therefore range profiles can be thought of as fundamentally one-dimensional signatures of radar targets as two of the three dimensions have little resolving capability. Their features rely on the geometry of the target, the particular Line-Of-Sight (LOS), the orientation of the target coordinate system with respect to a global system and the radar design parameters. When an adequate range resolution is achieved, a radar target may be resolved in a discrete group of scatterers in correspondence to those parts characterised by stronger electromagnetic returns. Therefore, scatterer energy reflections occur in different positions along the range profile and with different intensities, making available enough information about target geometry and orientation for classification purposes. The concept of resolution is often used with imprecise or varying meanings. Most often resolution is defined as the ability of radar (or any sensor) to distinguish between two closely spaced scatterers. A measure of this ability is captured in the radar system point spread function or impulse response function with resolving power being determined by the 3-dB points. This is a reasonable definition but care needs to be taken as there is an implicit assumption that the target to be considered has point-like properties only.

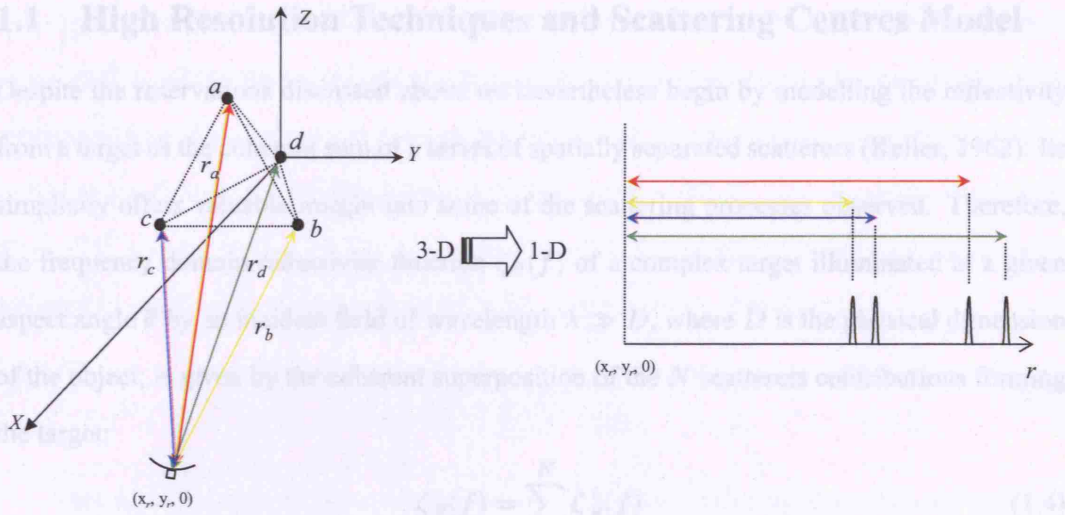

Figure 1.2: *Range profile from three scatterers* $(a, b, c$ *and* $d)$ *in a 3-D space: the distances* $(r_a, r_b, ..., r_d)$ *between the sensing system located in* $(x_r, y_r, 0)$ *and the scatterers are projected onto the 1-D range space.*

This is well known not to be the case but nevertheless has proved a useful descriptor of radar performance. As will be seen later if resolution is improved more scatterers on a target can be distinguished from one another and there is undoubtedly an improved amount of detail in the resulting signature. Care also has to be taken with synthetic aperture imaging in two dimensions such as Synthetic Aperture Radar (SAR) and Inverse SAR (ISAR). These imaging techniques again have an implicit assumption that targets have point-like properties and are continuously illuminated during formation of the image. Once again most targets cannot be modelled as a series of points and quite often there is occlusion of one target by another. For example the scatterers near the illuminated front of a target often place the scatterers further back in shadow and hence they are not imaged and thus the resolution in a typical SAR image is not necessarily constant, a fact often overlooked by developers of classification algorithms. However, these assumptions have still resulted in robust SAR and ISAR image generation and as larger apertures are synthesised there is an undeniable increase in detail in the resulting imagery. We will return to these themes as we explore high resolution techniques and approaches to classification of the resulting target signatures.

1.1 High Resolution Techniques and Scattering Centres Model

Despite the reservations discussed above we nevertheless begin by modelling the reflectivity from a target as the coherent sum of a series of spatially separated scatterers (Keller, 1962). Its simplicity offers valuable insight into some of the scattering processes observed. Therefore, the frequency domain reflectivity function $\zeta_\theta(f)$ of a complex target illuminated at a given aspect angle θ by an incident field of wavelength $\lambda \gg D$, where D is the physical dimension of the object, is given by the coherent superposition of the N scatterers contributions forming the target:

$$\zeta_\theta(f) = \sum_{i=1}^{N} \zeta_\theta^i(f) \tag{1.4}$$

where

$$\zeta_\theta^i(f) = A_\theta^i(f) \exp[-j\,\vartheta_\theta^i(f)] \tag{1.5}$$

Thus we see that the amplitude A and phase θ of the i-th scatterer depend on both frequency and aspect angle. In particular, when the Geometrical Theory of Diffraction (GTD) is considered, the model includes the half integer coefficient α, which depends on the target geometrical structure (Ross (1966), Plonus et al. (1978), Potter et al. (1995)):

$$\zeta_\theta^i(f) = A_\theta^i \left(j\frac{f}{f_c}\right)^{\alpha_i} \exp\left[-j2\pi\left(\frac{2fr_i}{c}\right)\right] \tag{1.6}$$

where A is now the complex amplitude of the i-th scattering centre, r_i its relative range, f_c the reference frequency. The model could be expanded introducing the length and orientation of the distributed scattering centres (Gerry et al., 1999). The Scattering Centers Model (SCM) for manmade targets, introduced by Rihaczek and Herschowitz (2000), has been validated (Miller et al. (2004)). The SCM considers only a small number of scatterers as features for ATR, each of them characterised by its position relative to the target structure and reradiation pattern which takes into account the fluctuations in amplitude and phase given by the particular observation angle. Furthermore, the Scattering Centre (SC) is identified by a peculiar scattering mechanism such as specular reflection, multiple bounces, surface waves, end-region scattering and cavity resonance.

An Inverse Fourier Transform (IFT) of 1.4 will yield the complex reflectivity function or range profile, where the magnitude in each of the IFT bins represents the magnitude of the reflections from the scatterers in a range resolution cell. In the simplest case, the resolution in the range dimension is related directly to the pulse width T_p (and therefore bandwidth $B = 1/T_p$) of the waveform. Two targets of equal RCS are said to be recognised as being resolved in range when they are separated from each other by a distance:

$$\Delta r = \frac{cT_p}{2} = \frac{c}{2B} \tag{1.7}$$

This equation tells us the range resolution of a pulsed radar system when the pulse is not modulated. This conventional definition has been referred to as "nominal" since it could be improved to the fundamental limit of the "functional" resolution c/B by considering the complex processor output instead of the intensity only (Rihaczek, 1996). Nevertheless a very short pulse is needed if High Range Resolution (HRR) is required. Indeed, to resolve all the scatterers unambiguously the pulse length has to be short enough such that only one scatterer appears in each range cell, notwithstanding potential superresolution and sidelobe suppression performance enhancements. Normally as high a range resolution as possible is used and it is accepted that not all scatterers will be resolved. The resulting ambiguity is one source of ensuing difficulty in the next stage of classification.

The concept of resolution is now explained using an example. We assume a transmitted bandwidth $B = 500$ MHz, and a single point scatterer located at two different positions: the first one is $r_1 = 30$ m and the second one is spaced by the nominal resolution (see equation 1.7) from the first position $r_2 = r_1 + \Delta r = 30.3$ m. If no weighting function is applied, the individual responses are shown in figure 1.3. It is worth noting that it is common to link the range resolution to the range from the -3 dB point of the point spread function, whereas it should be thought of as half the null-to-null width ($t_{n,n} = 2/B$) measured on the mainlobe that can be measured as 4 dB below the peak. Therefore, the -4 dB points of the two point spread functions intercept at the range resolution distance (30 cm). Therefore, in this case the capability of correctly locate the single point scatterer satisfy equation 1.7.

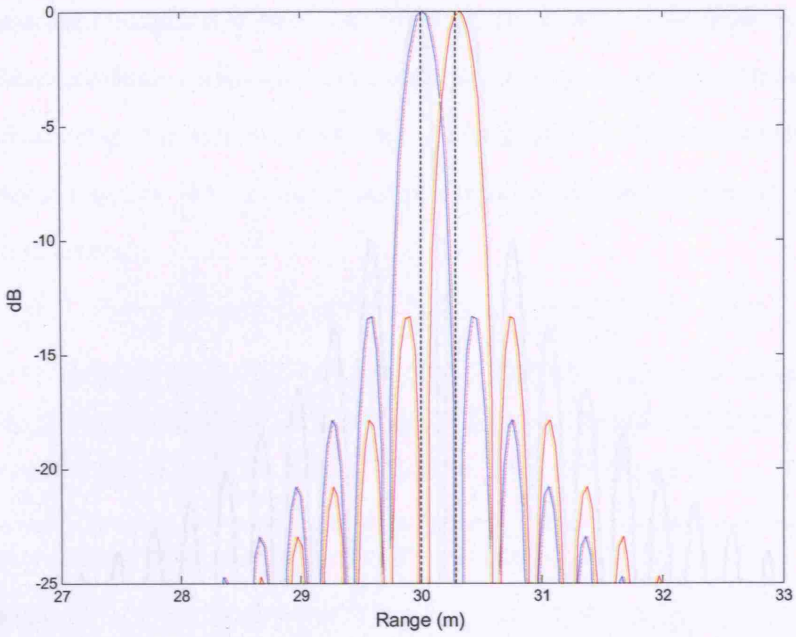

Figure 1.3: *Range location of a point scatterer at 30 cm range (blue) and 30.3 cm (red).*

If we consider a target formed by two point scatterers located in r_1 and r_2, their coherent interaction shows a different result, illustrated in figure 1.4.

The phase difference of the two scatterers is zero as a consequence of the point-scatterer assumption. As can be observed, the two points are resolved for a higher range than the nominal resolution. Nevertheless, their position is slightly shifted, as a consequence of the interference of main- and side-lobes. This phenomenon is more pronounced when the scatterers add an additional phase given by the particular structural geometry (the term $\alpha \neq 1$ in equation 1.6). In this case, their interference could be more accentuated depending on their distance with the result of deteriorating the positioning capabilities of the system although increasing their separation in range. If the complex target is also considered as a distribute set of scatterers, this aspect becomes even more severe.

Therefore, the nominal radar range resolution should be intended as the ability of positioning a single point-scatterer along the range direction. In addition, the higher the range resolution, the higher the possibility of resolving two scatterers.

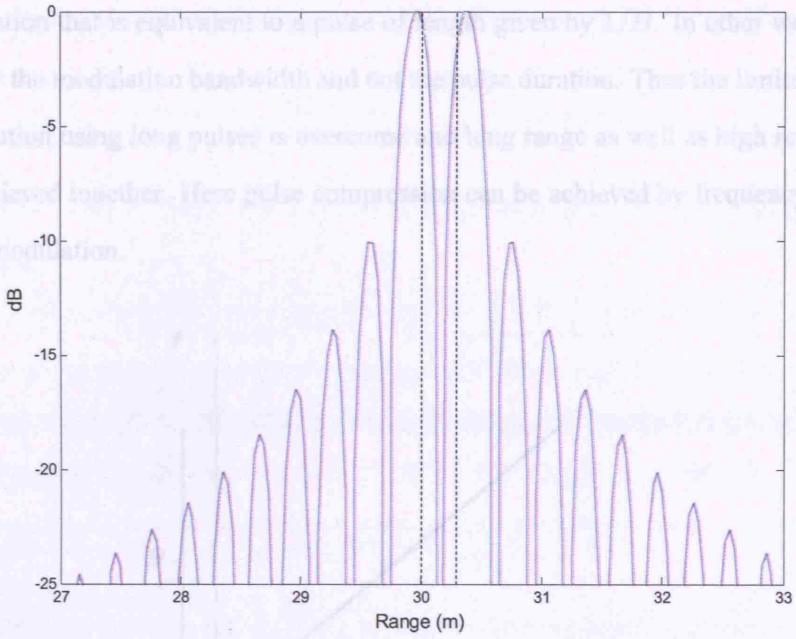

Figure 1.4: *Range profile of a two point scatterers target, spaced by the nominal range resolution (30 cm). The actual position is in correspondence to the two dotted lines.*

1.1.1 Pulse Compression

In long range radar, a long pulse is needed to ensure sufficient energy to detect small targets. However, according to equation 1.7, this long length waveform has poor range resolution. The use of a short duration pulse in a long range radar system implies that a very high peak power is required. There is a limitation on just how high the peak power of the short pulse can be. Ultimately the voltage required will 'arc' or breakdown risking damage to the circuitry and poor efficiency of transmission.

One method which is common place today that overcomes this limitation is pulse compression (Skolnik (1990), Sullivan (2004)). Pulse compression is a technique that consists of applying a modulation to a long pulse or waveform such that the bandwidth B of the modulation is greater than that of the un-modulated pulse (i.e. $1/T_p$). On reception the long pulse is processed by a matched filter to obtain the equivalent resolution of a short pulse (of width $1/B$). In the time domain the received signal is correlated with a time reversed replica of the transmitted signal (delayed to the chosen range). This compresses the waveform into a single

shorter duration that is equivalent to a pulse of length given by $1/B$. In other words it is determined by the modulation bandwidth and not the pulse duration. Thus the limitation of poor range resolution using long pulses is overcome and long range as well as high resolution can both be achieved together. Here pulse compression can be achieved by frequency, phase and amplitude modulation.

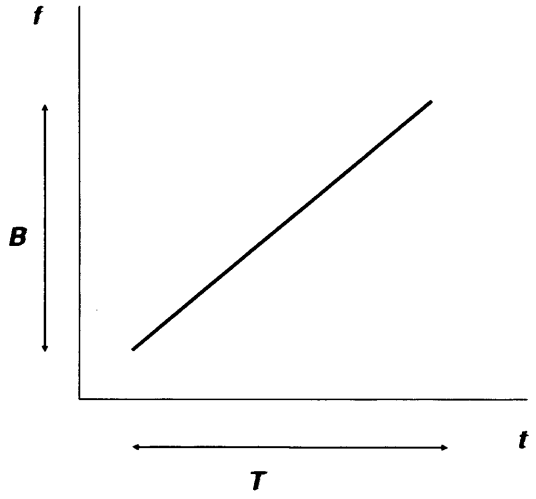

Figure 1.5: *Changing the frequency as a function of time across the duration of a pulse. B is the modulation bandwidth and T the pulse duration.*

The most common form of modulation used is the change the frequency from the start to the end of the pulse (figure 1.5) such that the required bandwidth B is swept out:

$$p(t) = \frac{1}{T_p} \text{rect}\left(\frac{t}{T_p}\right) \exp\left(\omega_c t + \pi\gamma t^2\right) \tag{1.8}$$

where ω_c is the carrier frequency and γ the rate of frequency variation with time. The instantaneous frequency is:

$$f(t) = \frac{1}{2\pi}\frac{\partial}{\partial t}\left(\omega_c t + \pi\gamma t^2\right) = \frac{\omega_c}{2\pi} + \gamma t \tag{1.9}$$

Therefore, the bandwidth of the frequency sweep is:

$$B_{sw} = f(t)\Big|_{T_p} = \gamma T_p \tag{1.10}$$

Since the signal reflected back by a point-scatterer is a delayed version of the transmitted signal, output of the matched filter is represented by the auto-correlation function, also known

as the Point Spread Function (PSF) $h(t)$:

$$h(t) = (T_p - |t|)\,\mathrm{sinc}\Big(\frac{\gamma}{\pi} t\,[T_p - |t|]\Big)\,\mathrm{rect}\Big(\frac{t}{T_p}\Big) \tag{1.11}$$

where sinc $x = (\sin \pi x)/(\pi x)$. The first null of $h(t)$ is often related to the temporal resolution of the signal r_t:

$$t = r_t = \frac{T_p}{2}\left(1 - \sqrt{1 - \frac{4}{T_p B_{sw}}}\right) \simeq \frac{1}{B_{sw}} \tag{1.12}$$

The approximation in 1.12 holds in the case of large time-bandwidth products and leads to a spatial resolution given by:

$$\Delta r = \frac{c r_t}{2} = \frac{c}{2B_{sw}} \tag{1.13}$$

The Pulse Compression Ratio (PCR) is equal to the time-bandwidth product $B_{sw}T_p$ and represents the range compression factor (i.e. the ratio of the pulse length before time compression to its length after compression $1/B_{sw}$). The PCR is a gauge of the range resolution improvement given by pulse modulation.

1.1.2 Stepped-Frequency Chirp

To achieve even higher range resolutions a frequency modulated Stepped-Frequency compressed waveform may be employed. This reduces the instantaneous modulation bandwidth requirement while increasing the overall bandwidth. In other words the necessary wider bandwidth waveform is synthesised using a number of individual pulses (figure 1.6). However, it should be noted that this has the disadvantage of collecting the individual waveforms over a finite period of time making the required coherency vulnerable to target motion. High Range Resolution HRR profiles or Synthetic Rnage Profiles (SRPs) are subsequently produced by processing a wideband reconstruction of a targets reflectivity spectrum in the frequency domain (Wilkinson et al. (1998), Wehner (1995)).

The Stepped-Frequency compressed pulses can be either recombined in the time (Wilkinson et al., 1998) or frequency domain (Nel et al., 2002). In figure 1.6, the frequency domain reflectivity reconstruction is shown schematically. A burst of N chirp-modulated pulses of bandwidth B_{sw} is transmitted. After collecting a sample version of each echo pulse, each

subspectrum needs to be translated in frequency by the amount:

$$f_i = \left(i + \frac{1-N}{2} \right) \Delta f \tag{1.14}$$

where Δf is the frequency step between pulses.

By adding the compressed individual portions of the reflectivity function - which results from time convolution between each received pulse with the complex conjugate of the corresponding transmitted pulse - the entire spectrum is eventually obtained commensurate with the extended bandwidth. The HRR profile may then be synthesised from an inverse FT applied to each row of the time history of the target frequency domain signature matrix. In order to avoid gaps in the synthesised reflectivity spectrum, the frequency step should be chosen at least equal to the chirp bandwidth. Furthermore, to avoid the nonlinearties of the side portions of the chirps spectra and subsequent possible artifacts, the frequency step is usually less to B_{sw}.

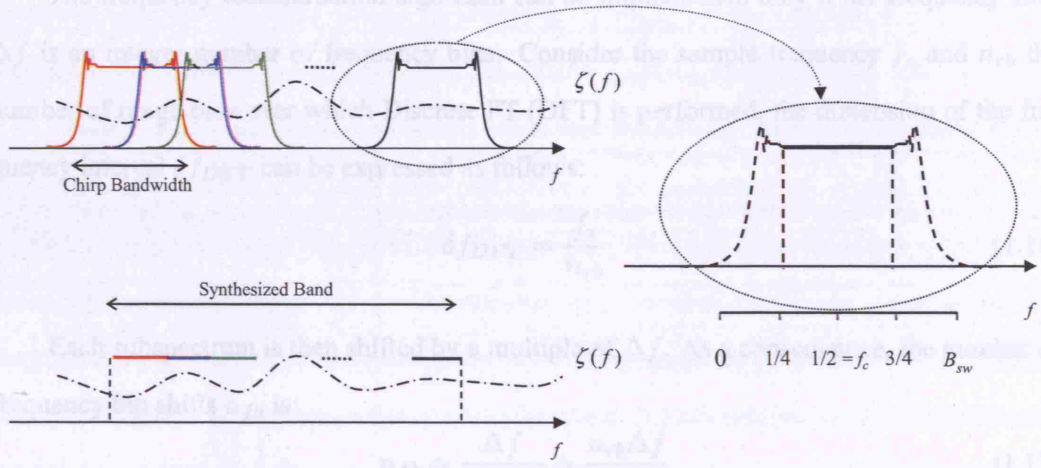

Figure 1.6: *Time representation of Stepped-Frequency waveform. The centre frequencies overlap over half of their bandwidth B_{sw}, so that the first and the last quarter of each sub-spectrum is discharged before the frequency domain reconstruction.*

As illustrated in figure 1.6, the total synthesised bandwidth B_{syn} relies on the choice of

both the frequency step Δf and the frequency sweep of each chirp B_{sw}:

$$B_{syn} = NB_{sw} - (N-2)\left[2\left(\frac{B_{sw}-\Delta f}{2}\right)\right] - 2\left(\frac{B_{sw}-\Delta f}{2}\right) \tag{1.15}$$

In 1.15, the first term represents the case in which subspectra don't overlap, that is a frequency step equal to the transmitted bandwidth. If $\Delta f < B_{sw}$, is the transmitted bandwidth, the second term takes in account the central subspectra: they overlap for the fraction $(B_{sw} - \Delta f)/2$, i.e. for each pulse that portion is discharged twice, once per overlapping side.

The first and last subspectra respectively overlap with the second and penultimate pulses. This accounts for the third term in the equation. If the inter-pulse frequency step increment is equal to $B_{sw}/2$, it is possible to synthesise a total bandwidth $B_{syn} = (N+1)B/2$. For example, considering a pulse duration $T = 300$ ns, a resolution of 45 m is obtainable without any further processing. If it is chirp-modulated with a $PCR = 75$, i.e. $B_{sw} = 250$ MHz, the range resolution improves until $\Delta r = 60$cm. Finally, using Stepped-Frequency Chirp techniques, involving $N = 6$ pulses stepped by $\Delta f = B_{sw}/2$, it is possible to achieve $B_{syn} =$ 875 MHz and, therefore, an improved range resolution $\Delta r \simeq 17$cm.

The frequency reconstruction algorithm can be implemented only if the frequency shift Δf is an integer number of frequency bins. Consider the sample frequency f_s and n_{rb} the number of range bins over which Discrete FT (DFT) is performed, the dimension of the frequency interval δf_{DFT} can be expressed as follows:

$$\delta f_{DFT} = \frac{f_s}{n_{rb}} \tag{1.16}$$

Each subspectrum is then shifted by a multiple of Δf. As a consequence, the number of frequency bin shifts n_{fb} is:

$$n_{fb} = \frac{\Delta f}{\delta f_{DFT}} = \frac{n_{rb}\Delta f}{f_s} \tag{1.17}$$

If n_{fb} is not an integer number, the coherent addition of subspectra yields superposition of their side parts in the combined spectrum. This problem affects the resultant high resolution range profile, generating sidelobes which are actually aliases of the main lobe.

An example of resolution improvements obtained using Stepped-Frequency Chirp modulation over a single compressed pulse is given in Chapter 3 (figure 3.7) where, after matched

filter operations, a series of High Resolution Range Profiles (HRRP) are compared.

1.2 Range Profile Variability

In a radar system, during the phases of collection and processing of data, a number of factors could affect range profiles, causing their variability although the orientation of the target and the system parameters remain the same (e.g. *measurement noise*, *rotational and translational range migration*). Other factors that may return signatures difficult to be correctly recognised are *speckle*, *multipath* and *shadowing phenomena*. We now consider these factors and their effect on the range profile in more detail.

Range profiles are affected by *measurement noise*. Electrical noise is electrical energy of random amplitude and frequency and it is present in every radio receivers above absolute zero. The thermal agitation in the input stage of receivers is amplified by the receiver full gain and becomes the main source of measurement noise. Other possible causes that could affect radar measurements are clutter (unwanted radar returns from scatterers), jamming (signals transmitted in order to deliberately confuse the system attempting to detect it), and Electro-Magnetic Interference (EMI) caused, for example, by other telecommunication systems in the vicinity of the radar.

Translational range migration occurs when the target is moving along the radar line of sight direction. In this particular scenario, since all the scatterers are shifted by the same number of range bins, the features of one-dimensional signatures from the target appear in the same position except for an offset. Translational range migration must be compensated out during any image reconstruction from sequences of range profiles, where a migrating scatterer appears in more than one range bin, causing blurring. Rotational motion of the target with respect to the line of sight over small angles (of the order of a few degrees) contributes to deforming of ISAR imaging. Additionally, if the motion is over greater apertures, it can yield *rotational range migration* of scatterers and give rise to both range profile variability and image blurring. This is indicated in figure 1.7.

Taking the example of an aircraft target and considering the wing span to be L metres, for the scatterers located on wingtips, a rotation of $\Delta\omega$ degrees produces a range migration of

$\sim L\Delta\omega$. In order to avoid rotational range migration, this migration has to be smaller than the range resolution of the radar system: $\sim L\Delta\omega < \Delta r$ (Hudson and Psaltis, 1993).

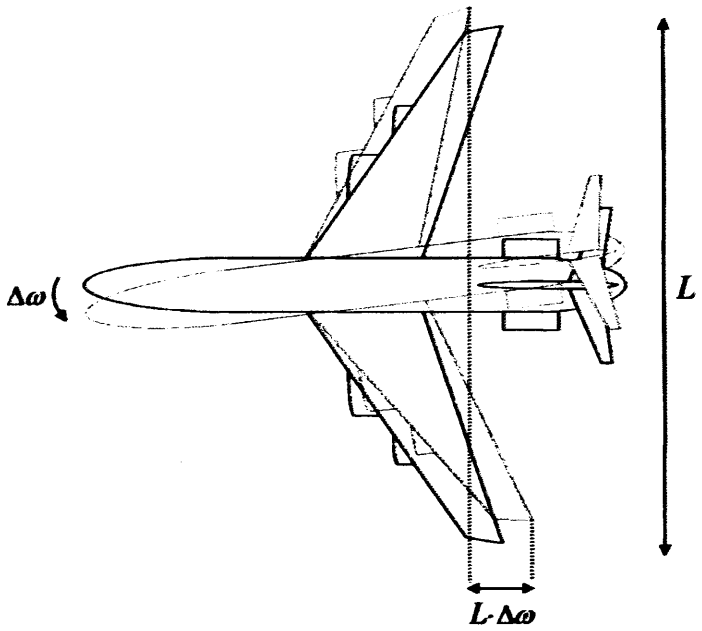

Figure 1.7: *Rotational Range Migration.*

In the presence of small translational or rotational movements of the target, that is when range migration of the scattering centre is avoided, the change of relative ranges of scatterers whose signatures go into the same range bin might lead to fluctuations of RCS as a consequence of interference effects. This phenomenon is known as *speckle*, always present in radar imagery. The aspect angle changes give rise to speckle fluctuations. When $L\Delta\omega \backsimeq \lambda/4$, where λ is the wavelength of the illuminating pulse, the scatterers at a distance L from each other can produce either complete destructive or complete constructive interference. In order to contrast range profile variability due to speckle fluctuation, the rotation of the target should be limited by small angles: $\Delta\omega \ll \lambda/4L$. For example, for an aircraft having a wing span $L = 10$ m illuminated at 9 GHz (X-Band), speckle fluctuations appear at any $\Delta\omega = 0.05°$. Therefore, characterising a target from all possible aspects implies a great number of range profiles. Furthermore, considering that the incident field is subject to bouncing reflection,

speckle fluctuation is considered inevitable, indicating a need to find approaches to reduce or account for this source of signature variability.

The aspect angle is relevant also for *shadowing phenomena*: for a given target heading, a number of potential scatterers may be occluded by other parts of the object (local shadowing) or by other targets (global shadowing) and, therefore, not be illuminated by the incident field. As a consequence, a portion of the target is unavoidably masked and the signature further corrupted. This may be thought of as a loss of information as it can often result in significant regions of the target not contributing to the measured signature or profile.

The interaction of the complex target with the surrounding scenario gives rise to another source of aspect dependent range profile variability known as *multipath*. If ATR is attempted, the delayed returns due to multiple reflection paths becomes problematic when the target is located in the proximity of highly reflective surfaces (ground and sea clutter) and the signal reaches the receiver covering different distances (figure 1.8): the direct ($2r$), the single bounce ($r + a + b$) and the double bounce paths ($2a + 2b$).

The sources of range profile variability have a similar influence when a set of signatures is collected in order to produce 2-D images of the target.

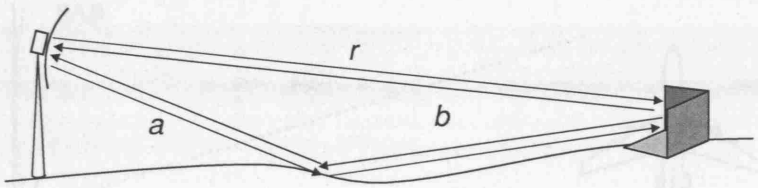

Figure 1.8: *Direct, single-bounce and double-bounce multipath.*

A number of techniques have been developed in order to isolate the clutter return from the target echoes, e.g. Space-Time Adaptive Processing (STAP; Klemm, 1999), Zero-Doppler Clutter Estimation and Subtraction (ZDC-ES; Showman et al., 1998). Other approaches have been investigating the speckle reduction, such as Polarimetric Whitening Filter (PWF; Novak

and Burl, 1990), Multi-Look PWF (MPWF; Liu et al., 1998) and measurement noise mitigation (e.g. Non-Coherent Averaging Skolnik (1990), Coherent Averaging Zyweck and Bogner (1995) and Target Normalisation Zyweck and Bogner (1996)).

It is clear that these factors can be overcome and averaged out by using further uncorrelated signatures from the target, that is using different perspectives. The employment of multiple perspectives is a way to reduce some of these effects through the exploitation of the additional information that exhibits aspect angle dependency.

1.3 Synthetic Aperture Radar

One-dimensional radar target signatures can be used to form two-dimensional radar images. Radar images are formed by coherent interaction of the transmitted microwave with the targets over an angular interval.

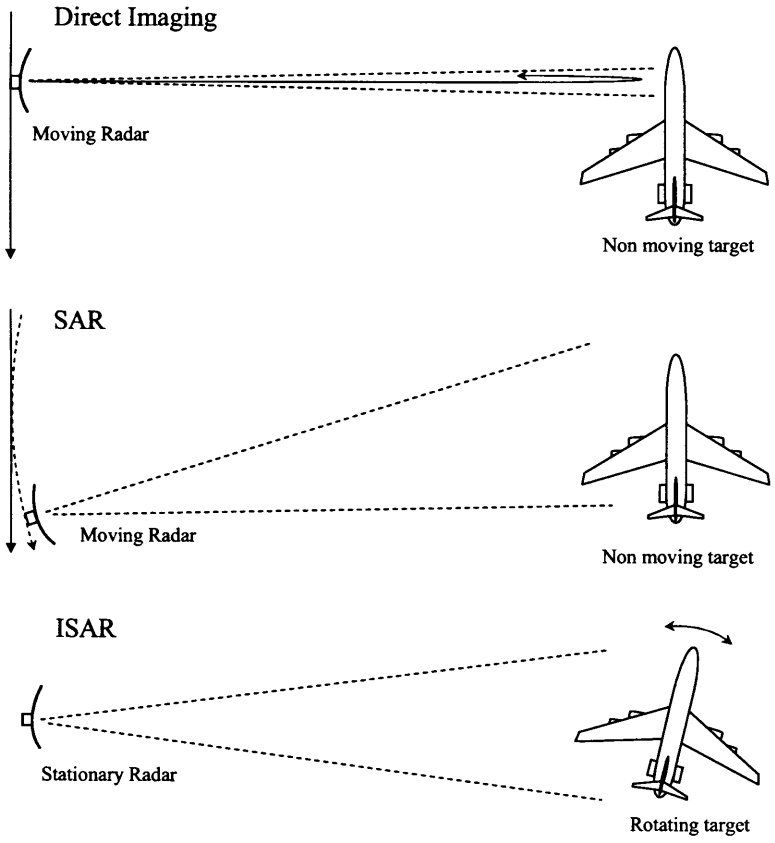

Figure 1.9: *Geometries of different imaging systems.*

The two-dimensional visualisation of radar target electromagnetic backscattering behaviour collected by a radar system is called a *radar image*. According to Mensa (1991), a radar image is a spatial distribution of reflectivity corresponding to an object. Two-dimensional imaging techniques offer a more detailed description of a target's scattering characteristics and are therefore expected to be a richer source of information for classification. Radar images are largely used for identifying and characterising the radar reflectivity function of complex targets. They provide the input to ATR systems as an identifier for classes of objects.

One method to obtain radar images is *Direct Imaging* (Figure 1.9), where a short-pulse radar scans the object in cross range direction. This technique, especially employed in sonar systems (side-scan sonar), requires a very narrow beamwidth to achieve high range resolution and the resolution deteriorates at long distances. An alternative one is the Synthetic Aperture imaging method. It uses not only a number of observations of complex targets from different perspectives but can also use frequency diversity to obtain improved down range high resolution.

SAR systems - as well as the corresponding Synthetic Aperture Sonars (SASs) - are used in order to obtain high resolution images of targets or maps of surfaces target areas and terrain (Curlander and McDonough, 1991).

SAR is fundamentally based on measurement from an effective long antenna whose dimensions are due not to the actual physical dimensions but to signal processing on the sensed reflectivity data (Cutrona, 1970). The term Synthetic Aperture is related to the path covered by the antenna during illumination of the target and its backscattered return data collection (Figure 1.10).

In this way it maps out a virtual antenna of very long dimensions thereby producing very high angular resolution.

Considering a real antenna with beamwidth α, the real beam sweeping over the target, the along-track antenna footprint length D at range R is:

$$D = R\alpha \tag{1.18}$$

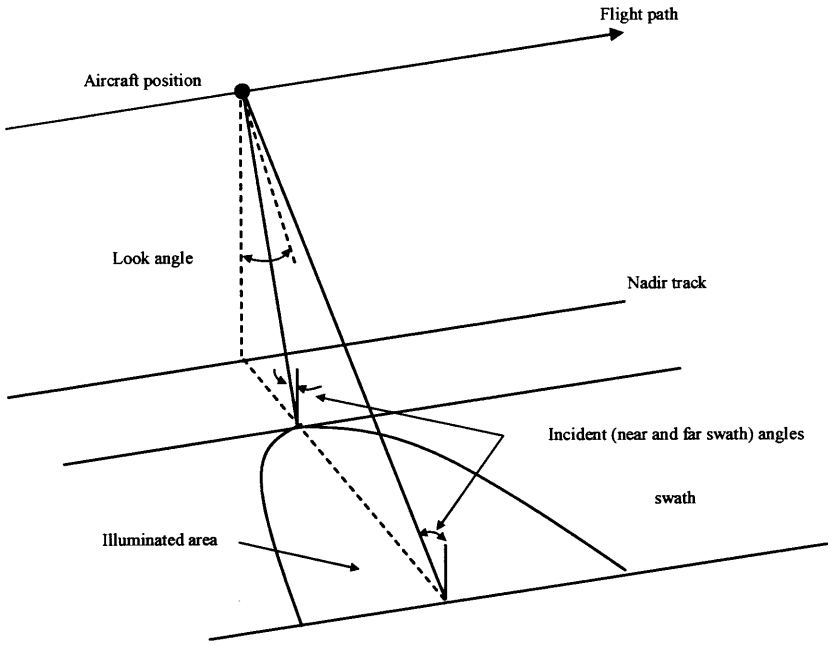

Figure 1.10: *SAR Geometry.*

The real antenna beamwidth is a function of the incident wavelength λ and the antenna length l:

$$\alpha = \frac{\lambda}{l} \tag{1.19}$$

For a synthetic antenna, the beamwidth α_s can be written as:

$$\alpha_s = \frac{\lambda}{2D} = \frac{l}{2R} \tag{1.20}$$

Comparing α and α_s, the angular resolution of the SAR is improved by a factor two. This is due to the two-way propagation of the signal before the synthetic antenna is formed. Furthermore, the beamwidth resolution of a SAR is inversely proportional to the range R. The cross-range resolution Δr_c for a SAR system is then:

$$\Delta r_c = R\alpha_s = \frac{l}{2} \tag{1.21}$$

As a conclusion, the along-track maximum resolution in side-looking SAR systems is independent of the range of the scatterer from the radar and the wavelength of the incident waveform.

1.4 Inverse Synthetic Aperture Radar

As can be seen from figure 1.9, an ISAR geometry is directly related to the SAR geometry. In SAR scenarios the target is still and the system moves around it. In ISAR the radar is still and it collects information during the time the moving target is inside its illuminating beam. In ISAR systems, the slant range resolution can be achieved by transmitting wideband pulses as explained in Section 1.1, whilst cross range resolution is linked to the angular aperture synthesised during the rotation of the target. Of course hybrid imaging involving platform and target motion is also possible.

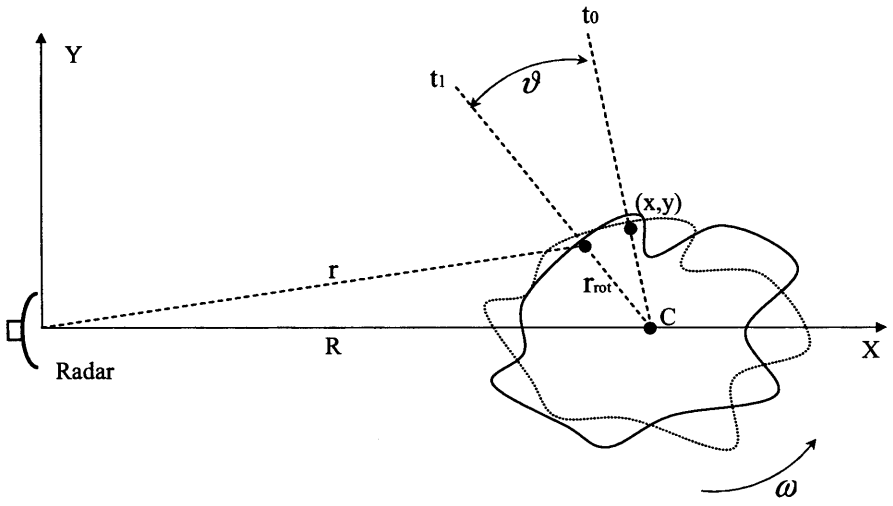

Figure 1.11: *Range for a Generic Scatterer in ISAR Geometry.*

The angular velocity of each scatterer of the target is related to its range from the radar. In figure 1.11, the target is rotating with angular velocity ω around a centre of rotation C, located in $(R, 0)$.

The radar is located on the centre of a system of coordinate and a scatterer in (x, y). From time t_0 to the time t_1, the point scatterer rotates for an angle $\theta = \omega(t_1 - t_0)$. As a consequence,

its distance from the centre of coordinate R is expressed as follows:

$$r = \sqrt{(R\sin(\theta) - x)^2 + (R\cos(\theta) - y)^2} \tag{1.22}$$

The distance r from the radar is the actual range of the scatterer and, during the rotation, the sequence of range bins covered by the set of scatterers forming the target build up the necessary information in order to form the final two-dimensional DFT (2-D DFT) ISAR image.

In ISAR image processing, the motion of the target is unknown. The target motion can be seen as the superposition of rotational and translational motions with respect to the radar system. If the former contributes to the ability to resolve in cross-range, in order to obtain a focused image of the target it is necessary to compensate for phase errors due to the translational motion occurring during data collection. This is usually referred to as *motion compensation*, whereas further residual motions may require correction using *autofocus techniques*, and apply to SAR processing as well. After this correction the image may be obtained by processing data collected over an arc of circular aperture whose dimensions depend on the rotational speed of the target and the integration time. Doppler shifts of each scatterer of the target are proportional to the instantaneous angular velocity. Assuming the target is only rotating without any translational movement, i.e. there are no radial velocity components for the point C (figure 1.11) along the line of sight, and that the target is rotating with constant angular velocity ω in rad/s about an axis perpendicular to X, its Doppler frequency is given by (Wehner, 1995):

$$f_D = \frac{2\omega r_c f_c}{c} \tag{1.23}$$

where r_c is the projection of the distance between the scatterer (cross range) and the centre of rotation on the direction $\bar{Y}$, and f_c the centre frequency of the incident waveform. The product $\omega \cdot r_c$ takes into account the instant relative velocity with respect to the centre of coordinate system. Considering two scatterers located in the same range bin, i.e. the same distance to the radar, but with different cross range from the centre of rotation $\delta r_c = r_{c1} - r_{c2}$, using 1.23, in the final image they are spatially resolved thanks to their different Doppler shift δf_D:

$$\delta r_c = \frac{c}{2\omega f_c}\delta f_D \tag{1.24}$$

Ultimately, the Doppler frequency resolution Δf_D determines the ability of the system to distinguish returns with different Doppler shifts, $\Delta f_D \approx 1/T$. This means that the Doppler resolution is proportional to the inverse of the available coherent integration time T. In a stepped-frequency chirp technique, the integration time is expressed as the time length of a series of bursts of pulses. As a consequence, the cross range resolution Δr_c, defined as the ability to resolve point scatterers along the cross range direction, is expressed as follows (Wehner, 1995):

$$\Delta r_c = \frac{c}{2 f_c \omega T} = \frac{1}{2} \frac{\lambda}{\Delta\theta} \tag{1.25}$$

Equation 1.25 exhibits the link between cross range resolution, wavelength λ of the incident field and rotation angle $\Delta\theta$ covered by the scatterers during the integration time T.

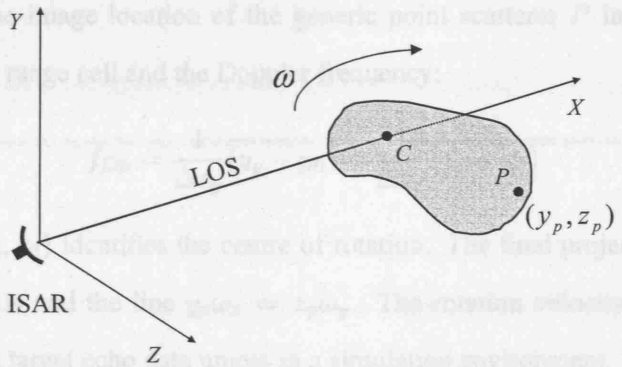

Figure 1.12: *Geometry of a target rotating around the point C.*

In order to avoid problems due to range migration and subsequent spreading of energy from a single scatterer into more range bins, a small arc of aperture must be processed, yielding a partial image of the target.

A radar image of an object is a spatial distribution of three-dimensional scatterer reflectivity projected onto a plane whose parameters are determined by the line of sight of the system and rotational components of the target (the *projection plane*).

If we consider the geometry shown in figure 1.12 for a point P at (y_p, z_p), where the radar is located in the origin of a coordinate system whose axis X coincides with the line of sight between the radar and the reference centre of target's rotation C, defining ω_y and ω_z as

the two components corresponding to velocity vectors on the plane whose axis is respectively Y and Z then Genyuan et al. (2001):

$$\omega_0 = \sqrt{\omega_y^2 + \omega_z^2} \tag{1.26}$$

The rotational motion ω_0 does not take into account rotational motions perpendicular to the line of sight ω_x because of their irrelevance for Doppler shifts. Therefore, only Y and Z components of the rotational motion result usable for resolving in Doppler frequencies. From equation 1.25 it is possible to write the two components Δr_y and Δr_z of resolution:

$$\Delta r_y = \lambda/2\omega_y T \tag{1.27}$$

$$\Delta r_z = \lambda/2\omega_z T \tag{1.28}$$

As a result, the image location of the generic point scatterer P in the radar image is defined by the slant range cell and the Doppler frequency:

$$f_{Dp} = \frac{1}{\Delta r_y}(y_p - y_0) + \frac{1}{\Delta r_z}(z_p - z_0) \tag{1.29}$$

where the point (y_0, z_0) identifies the centre of rotation. The final projection plane is determined by the X axis and the line $y_p\omega_z = z_p\omega_y$. The rotation velocity ω_0 of the target is not known from the target echo data unless in a simulation environment. Therefore, the cross range bin size and image plane are usually said to be unknown.

To obtain high resolution range profiles for ISAR, frequency compressed pulses are transmitted and received by the system. The target range profile is then produced at the output of the pulse compression filter at an intermediate frequency (as previously explained). If Stepped-Frequency processing is used, a spectrum reconstruction follows the quadrature detection, the sampling and digitising procedures. From each burst a range profile is ultimately obtained and, as a result of two-Dimensional DFT (2-D DFT) processing, it yields a partial image of the target for small angle of integration. In figure 1.13, the image processing sequence for a generic ISAR system is shown for a sequence of N bursts.

Data are collected in the Time-Domain in form of a sequence of in-phase and quadrature components $(I + jQ)'$. A number of n_s samples identifies a single chirp $h(k, m)$. Therefore

$M = n_s \cdot n_c$ samples are processed for each range profile, where n_c is the number of chirps for a burst. After the frequency domain spectrum reconstruction, a time history of frequency $(I + jQ)$ samples of $H(k, i)$ for the k-th burst and i-th step is obtained.

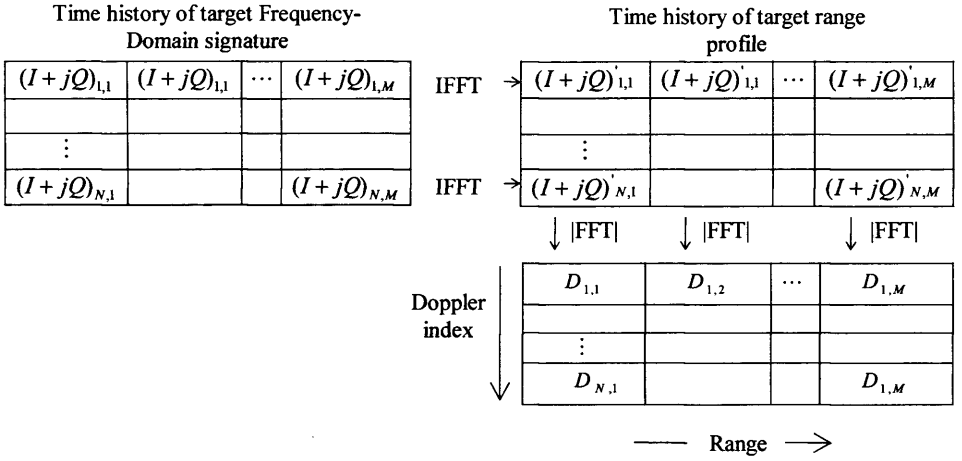

Figure 1.13: *Image processing sequence for an ISAR system using stepped-frequency chirp technique.*

The instrument to obtain the final radar image from the time history of frequency domain target signatures is the 2-D DFT. A matrix of magnitudes $D_{m,n}$ in m-th range cell and n-th Doppler cell of range-Doppler image is then produced. The two-dimensional DFT involves a sequence of one-dimensional DFT, that means it is achieved by first inverse transforming each row of a matrix:

$$h(k, m) = \frac{1}{M} \sum_{i=1}^{M} H(k, i) \cdot e^{j2\pi mi/M} \tag{1.30}$$

Each row is then replaced with its Fourier transform. Subsequently each N elements column is transformed and eventually replaced with its transform:

$$D(m, n) = \frac{1}{N} \sum_{k=1}^{N} h(k, m) \cdot e^{-j2\pi nk/N} \tag{1.31}$$

By combining equations 1.30 and 1.31, this procedure can be done in one single step and it is known as 2-D DFT:

$$D(m, n) = \frac{1}{MN} \sum_{k=1}^{N} \sum_{i=1}^{M} H(k, i) \cdot e^{j2\pi(mi/M - nk/N)} \tag{1.32}$$

By plotting this matrix, the final radar image for a burst sequence of N pulses, which corresponds to a certain synthesised aperture θ, is ultimately produced.

1.5 Radar Image Reconstruction

Rihaczek and Herschowitz (2000) consider the range resolution as the primary requirement for target identification, while the function cross-range resolution is to "aid" the resolution along the more significant down range direction. This is because of the relative simplicity of achieving high range resolution in contrast with the cross range resolution dependency on the target motion parameters. Although the two trihedrals in figure 3.9 can be distinguished in the two dimensional projection plane even at the lowest cross range resolution, their actual position - a crucial aspect for classification - can be accurately determined only by synthesising large apertures. A possible alternative is the image reconstruction combining multiple views of the target. A necessary step before Multi-Look or Tomographic image reconstruction is to establish the relative centre of rotation of the target in order to exactly superimpose the pixels related to moving scatterers.

1.5.1 Multi-Look Processing

Multi-Look image reconstruction is a technique that allows production of radar images of rotating targets by synthesising wide angles. A limit of 2-D DFT to perform complete target images is range bin migration: when the rotation angle becomes greater than a certain limit, a point scatterer changes range bin, its Doppler information deteriorates and the scatterer's energy goes in other range bins causing a blurred image.

Range bin migration can be overcome by processing data collected from small arcs of a circular aperture, taking a number of images, rotating them by the corresponding angle covered and then superimposed. This method is known as Multi-Look Image Processing or Enhanced Image Processing (Ausherman et al., 1984). As shown in figure 1.14, partial images are taken synthesising small apertures θ_{ap} so that the small-angle approximation is valid and range migration does not need to be performed.

The illuminating direction is the left-hand side of the image, so that the dashed outline is locally obscured. The corresponding intensity matrices are subsequently rotated of the same

angle by the rotation matrix R_{ap_n} expressed by:

$$R_{ap_n} = \begin{bmatrix} \cos(\theta_{ap_n}) & \sin(\theta_{ap_n}) \\ -\sin(\theta_{ap_n}) & \cos(\theta_{ap_n}) \end{bmatrix} \tag{1.33}$$

The partial images are non-coherently combined to form the final reconstructed image. This procedure allows the recombination of different perspectives and therefore the detection of the features that are occluded when the target is sensed from a particular direction.

The non-coherent averaging also guarantees speckle, measurement noise and multipath reduction.

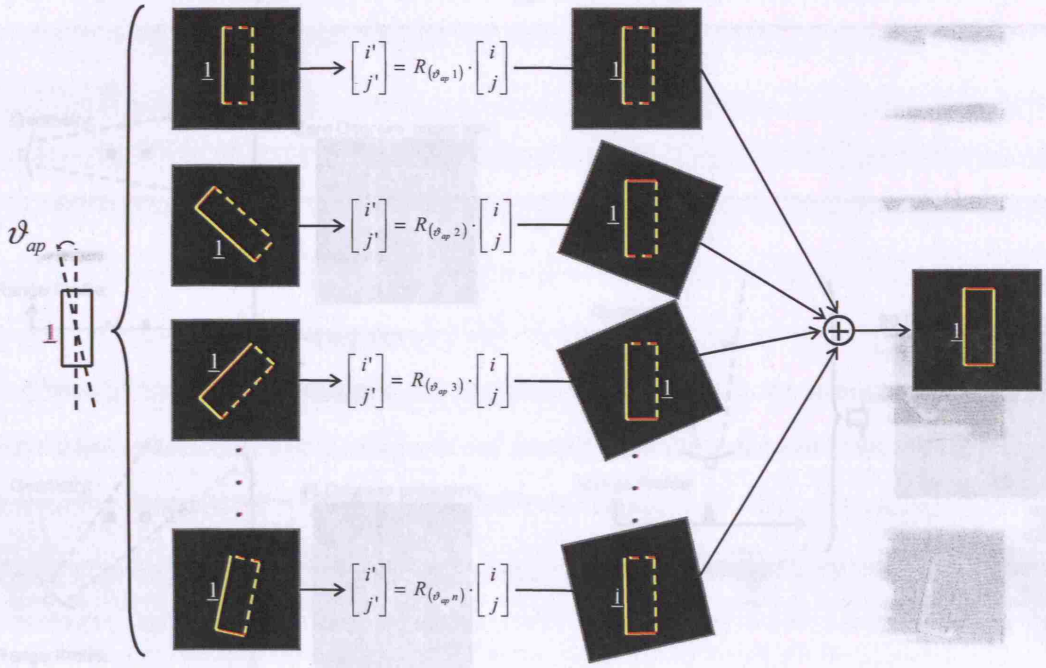

Figure 1.14: *Enhanced Image Processing.*

1.5.2 Tomography

Considering a SAR data collection geometry, in particular the sensor travelling along a circular path around the target, the line of sight vector is now a line from the radar to the centre of rotation. For each burst $l = 1...L$ a range profile is taken, that means the radar line of sight slices through the scene collecting data from aspect angle θ_l. The Tomographic Reconstruction (TR)

algorithm (Kak and Slaney (1988), Munson et al. (1983)) applies the projection-slice theorem of the Fourier transform to compute the image which sates that the Fourier transform of a one-dimensional projection of a two-dimensional function $f(x, y)$ made at angle θ_l is a slice through the two-dimensional Fourier transform of $f(x, y)$ at angle θ_l. In figure 1.15, a generic scenario with two stationary targets is shown. They correspond to an ISAR geometry with two rotating targets. Whereas the Multi-Look 2-D DFT algorithm previously discussed converts the signals from many bursts simultaneously into a reflectivity image using 2-D DFT, TR generates an image by projecting each range profile individually back onto a two-dimensional grid of image pixels (white dashed lines in figure 1.15). This operation is referred to as *Filtered Back-Projection* and represents an inversion algorithm for the Radon Transform.

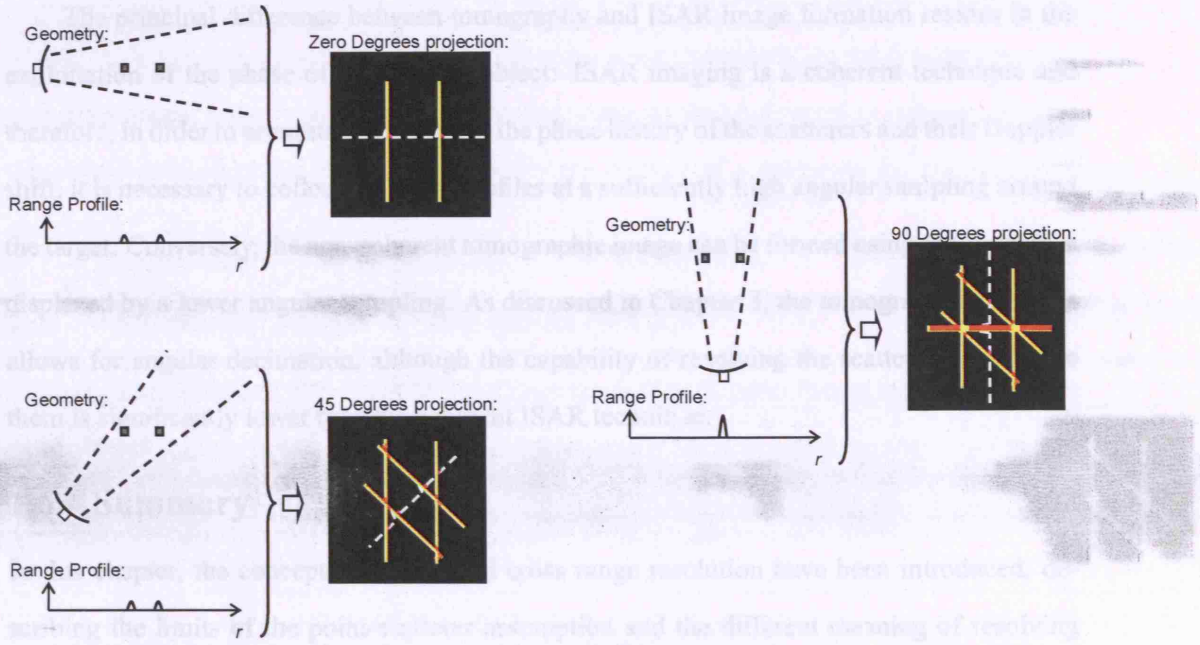

Figure 1.15: *Tomographic radar image reconstruction from a sequence of range profiles: geometries and projections.*

Then, lines perpendicular to the line of sight are drawn with intensity proportional to the magnitude of the corresponding range bin's magnitude. A single radar return resolves scatterers in range only so, as can be seen from the 90° projection, many scatterers can occupy the

same range bin. In that case, the magnitude is greater and the line drawn is more marked. Ultimately, by adding together pixel values, the intersections underline the location of scatterers and the final tomographic image.

The tomographic reconstruction technique carries a high computational burden which also limits its use but it allows a partial reconstruction of radar image (useful for automatic target detection) and it does not require a small step of rotation between consecutive profiles of the target that means the need for less information than other imaging algorithms to have rough representations of targets. Another favorable attribute of TR is that it maps well to a parallel processing architecture because the reconstruction operation is made on a pixel-by-pixel basis so that the image may be divided in a number of sub-images and computed separately by different processors and eventually put together (Desai and Jenkins, 1992).

The principal difference between tomography and ISAR image formation resides in the exploitation of the phase of the moving object: ISAR imaging is a coherent technique and therefore, in order to accurately reconstruct the phase history of the scatterers and their Doppler shift, it is necessary to collect the range profiles at a sufficiently high angular sampling around the target. Conversely, the non-coherent tomographic image can be formed using range profiles displaced by a lower angular sampling. As discussed in Chapter 3, the tomographic approach allows for angular decimation, although the capability of resolving the scatterers and locate them is significantly lower than the coherent ISAR technique.

1.6 Summary

In this chapter, the concepts of range and cross range resolution have been introduced, describing the limits of the point-scatterer assumption and the different meaning of resolving and locating different scattering centres. The main radar imaging techniques have also been analysed, both in one and two dimensions. The combination of target signatures using multiple perspectives has also been described using Multi-Look processing and Tomography.

The result of the techniques and definitions given in this chapter are used to produce and describe the patterns to be recognised as explained in Chapter 2. Subsequently, this concepts are applied on real target measurements and discussed in Chapter 3.

Chapter 2

Pattern Classification and Automatic Target Recognition (ATR)

In this chapter, the ATR literature survey reviewed in the introduction is expanded with its historical evolution and a description of the main processing techniques involved using one- and two-dimensional radar imagery, focusing the attention on their information content. Finally, the classification problem and its terminology are defined as a reference for the following chapters.

2.1 Automatic Target Recognition (ATR)

In military applications, Cooperative techniques (e.g. Identification Friend or Foe (IFF)) are combined with other sources of signature recognition in order to increase the overall reliability of possible engagement decisions (Automatic Target Recognition (ATR), also referred to as Non-Cooperative Target Recognition (NCTR)).

In this section, after introducing the historical IFF and NCTR background, the pre-processing techniques are described along with feature extraction methods.

2.1.1 The Historical Evolution of IFF/NCTR

The idea of remotely discriminating objects in terms of their hostility has its roots during the World War II, when the swiftness of aircrafts stated the problem of visual identification limits, historically accomplished using uniforms, flags and banners. Aircraft recognition based on observations, pioneered by Saville-Sneath (1941), can be considered the first attempt of

NCTR. Similarly to radar ATR, it was highly dependent on the target orientation and the main features analysed by human operators were the fuselage, the tail unit and the wings, nowadays representing the main scattering centres used for radar classification. With the advent of radar systems, the question of identifying the detected aircraft became crucial. For a long time, British radar operators were confused by the manoeuvres of German aircraft until they realised that the behaviour followed a signal from the ground. The Luftwaffe ground crew would transmit the signal to request a manoeuvre that caused a change in the received radar signal: the pilots rolled their aircraft to induce a detectable change in the signal polarisation that was an early form of electronic IFF. Modern IFF systems are Question/Answer systems (MK YY transponder systems and the Precise Participant Location and Identification (PPLI)), based on cooperative response to interrogations. They are not satisfactorily trustworthy because of their weakness of being jammed, replicated or damaged in combat environments. Furthermore, they cannot provide positive friend identification in dense environments since the interrogations are not "addressed". Although Other cooperative techniques such as Rules of Operation (e.g. following particular flight corridors or kinematics appropriately set for the mission) are often integrated with IFF, the All Service Combat Identification Evaluation Team (ASCIET), an American program designed to assess and improve Combat Identification (CID), still relies on Visual Identification (VID) before declaring the ID as hostile (Defence Science Board Task Force (1996), ASCIET (2000)). Other approaches need to be combined with cooperative systems such as NCTR techniques that utilise other sensors (e.g. radar, sonar, infrared and optical sensors) to improve the classification and reliability performance. Their signatures are selected and combined on the basis of range and weather conditions constraints. The utility of NCTR is well established in terms of range extension, fratricide reduction, collateral damage reduction and kill effectiveness. Nevertheless, the NCTR algorithms of current military systems are not robust enough to guarantee effective radar recognition capabilities.

2.1.2 NCTR: Process and Techniques

Figure 2.1 shows the main stages of radar ATR: within the received signal preprocessing, the target *detection* is performed by isolating a possible target's echo from the measurement noise;

the progressive reduction of the Region of Interest (ROI) is subsequently approached using *discrimination* techniques in order to distinguish the target backscattering from clutter; after passing targets of interest only, filtered by the *preclassification*, the signal is then processed by the following feature extraction stage. Finally, the classifier computes the membership decision according to the particular algorithm implemented. This will be the subject of the next section.

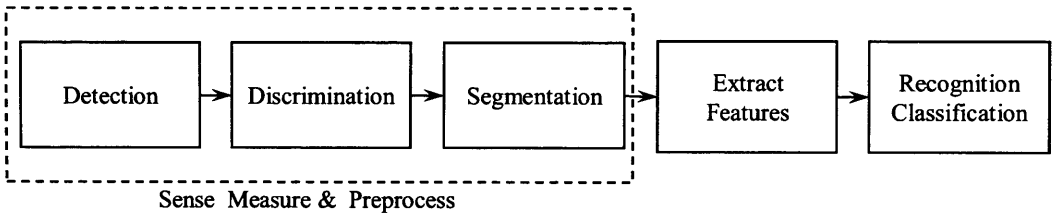

Figure 2.1: *NCTR stages.*

NCTR techniques can involve active (^{a} – the system processes the received version of the transmitted signal after being modulated by the target) and/or passive (^{p} – the system "listens" analysing signals generated either by the target or by an illuminator of opportunity and reflected by the target) sensors.

A possible taxonomy of the NCTR approaches based on different sensors measurements can be summarised as follows:

- **Radar**:
  - RCS Measurements $^{a/p}$;
  - Time-Frequency analysis $^{a/p}$;
  - High Range Resolution range profiles $^{a/p}$;
  - Synthetic Aperture Radar (SAR) or Inverse SAR (ISAR) images $^{a/p}$;
  - Multi-frame data $^{a/p}$ (i.e. the time evolution of all the above);
- **Sonar**:
  - Sidescan echoes ^{a};

- Synthetic Aperture Sonar (SAS) or Sidescan images [a];
- Multi-frame data [a].

- **Electronic Support Measures (ESM)** [p] (also known as Radar Warning Reciever (RWR) e.g. the estimation of the emitter parameters such as pulsewidth, PRF, beamwidth and rotation rate);

- **Electro-Optics Infrared (EO/IR)** (e.g. Infrared Search and Track (IRST), Forward Looking Infrared (FLIR))[p];

- **SIGnals INTelligence (SIGINT)** [p] (e.g. Communication Intelligence (COMINT), a technique that analyses the source and content of over-air transmitted messages).

Radar signatures can be combined with fully polarimetric techniques in order to further characterise the target backscattering.

Sonar systems can either be used underwater or in air. The techniques used to form images are similar to the radar ones although the absence of polarimetry and different backscattering properties and ensonifying frequencies.

SIGINT algorithms, described by Moir and Seabridge (2006) and ESM measures, diffusely treated by Neri (2006), can be thought of as other NCTR techniques useful when integrated with other sensed attributes from the target since it is not always possible to capture communication or radar signals from the targets and, furthermore, they would be ambiguous with the increase of the number of targets in the battlefield (Cohen, 2006).

Different problems arise from EO/IR techniques in relation to their ranging capabilities: although a higher wavelength is used and higher angular resolution achieved, these systems suffer atmospheric attenuation and unfavourable weather conditions.

We now concentrate our attention on radar based NCTR techniques using high range resolution profiles (one-dimensional) and SAR/ISAR images (two-dimensional) as possible techniques that could be combined with the mentioned ones above.

2.1.2.1 One-Dimensional Profile Classification

Radar targets can be identified either by forming images from a sequence of range profiles with enough resolution that they can be recognised by human observers, or by forming representations of the target for automatic target recognition. These time domain representations (as previously discussed in Chapter 1) can be either a one-dimensional projection along the line-of-sight or a two-dimensional image of the target. One-dimensional and two-dimensional data characterisations of targets need to be pre-processed before the classification task in order to eliminate redundancy in the data and isolate and distinguish useful information.

Pre-Processing Techniques. Some of the pre-processing techniques to reduce the noise level and make the target clear in the range window of interest are clutter and Zero-Doppler Clutter (ZDC) subtraction, Non-Coherent Averaging, Coherent Averaging, Target Normalisation and Target Alignment and Localisation.

Clutter and ZDC Subtraction: When stationary radars collect profiles from moving targets surrounded by clutter, the data exhibit a contribution from ZDC, stationary or quasi-stationary clutter that contributes to the HRR profiles. By observing a sequence of range profiles it becomes clear how the ZDC affects the classification process. Grass and trees show a relatively fast fluctuating term whose effect can be observed over relatively small turntable rotation angles. This is in contrast to the more persistent response given by spatially more restricted scatterers such as static ground discontinuities and mounds of gravel. This also means that the ZDC can be very different from one target measurement to another since as the vehicle approaches the turntable it modifies the local clutter structure and hence the ZDC component of the return in the signature. As a consequence, if ATR is implemented using the same data for both training and testing the classifier, the overall accuracy could be significantly upset by the ZDC contribution to backscattering, as the classifier would otherwise use the clutter as an input. Thus the ZDC has to be estimated and subtracted. This can be performed using data derived in the frequency domain following the method outlined by Showman et al. (1998). Conversely, when an airborne system collects range profiles from targets on the ground, unless they are moving (in which case GMTI algorithms can be applied to isolate them from the

clutter), the discrimination problem becomes more difficult and needs to be undertaken using statistical analysis of the clutter.

Non-Coherent Averaging: a single range profile may be affected by low signal-to-noise ratio (SNR), thus making classification difficult. In order to increase the SNR of signatures, it is possible to average a sequence of range profiles. The main consequence, as well as increasing the SNR, is the reduction of speckle. As demonstrated by Skolnik (1990) a low SNR can be improved by a factor $\sqrt{N}$ as a result of non-coherently averaging a sequence of N range profiles. If this pre-processing technique is used, it is necessary to align the range profiles.

Coherent Averaging: this technique yields higher SNR improvements when compared with the Non-Coherent Averaging (Zyweck and Bogner, 1995). On the other hand, this processing is particularly sensitive to the scatter Doppler shift since the coherent summation would lead to lower SNR if the phase is not properly compensated.

Target Normalisation: in order to make the target signature more prominent in the range window of interest, this procedure isolates the non-target return and decreases the noise level in the same region. This is achieved firstly by recognising the target region by measuring the window average magnitude level η_w. Then, the target region is found, which corresponds to values greater than η_w. Along the non-target region, the noise mean η_n is subsequently measured and finally subtracted from the non-target area (Zyweck and Bogner, 1996).

Target Alignment and Localisation: before non-coherent averaging of a number of range profiles, they need to be aligned, i.e. translated along the range direction allowing the complex target falling in the same range bins as much as rotational motion concedes. This is usually made by cross correlation, which works well because the change in aspect angle is assumed small. Localisation is related to the position of the target inside the range window. This is an important task for the next step of feature extraction. For example, if the first peak position is taken as a feature and localisation is not made, the final classification performance would deteriorate.

If position invariant signatures (such as the overall magnitude of the range profiles or frequency domain signatures) are taken as data for classification, the target does not need to be

located or aligned. Furthermore, in case of turntable data, no alignment is necessary because all the range profiles are aligned with respect to the centre of the turntable and the scatterers of interest fall in a range window centred on it.

Feature Extraction. Radar target classification can be carried out by using feature vectors extracted from one-dimensional signatures from targets such as range profiles. As previously discussed, a range profile can be seen as the projection of the apparent target scattering centres onto the range axis. High resolution range profiles provide a large amount of information which can aid recognition. On the other hand, the dimensionality of the data-set increases and it becomes more difficult to implement an accurate classifier. This problem is known as the curse of *increasing dimensionality* and it is discussed by Duda et al. (2001). Although a degradation of global information, after dimensional reducing the feature space, the problem often becomes more tractable. Feature extraction is by definition a form of dimensional reduction. On range profiles, it is possible to reduce the redundancy of range bin values by extracting essential information. This can be explained by the point-scatter return, for instance, which is projected onto a number of range bins shaped by the point spread function. Nevertheless, the peak value and position are the only essential parameters to characterise this scenario.

A possible set of features extracted from range profiles can be:

- The RCS of the target measured at different frequencies.
- The relative distance between the first peak and the last one detected. This information is strictly linked to the projected length of the target from the particular aspect angle it is illuminated.
- The k dominant peaks between the first and the last peak. Their intensities and relative distances are related with information on other feature of the target such as its position and orientation.

Peaks in the range profile can be found when the derivative of the signature is equal to zero. The peak size can be taken to be the width where the amplitude is half of the peak amplitude. The peak position, width and amplitude are $3 \cdot k$ features and can be used as an input for the

classifier (Purnell et al., 1998). The main drawback of this scattering model is the coherent interactions between the sidelobes of scatterers located one next to the other which could yield constructive or destructive interference, making it difficult to understand the number and location of the actual scattering centres. Furthermore, the dynamic range inside the profile is reduced since the low level scatterers are unlikely to be selected and therefore are discarded. This methodology can be refined using superresolution techniques such as the CLEAN algorithm (Tsao and Steinberg, 1988).

2.1.2.2 Two-Dimensional Image Classification

Radar images (as previously stated) are a two-dimensional representations of the target projected onto the projection plane identified by the line-of-sight from the radar to the centre of rotation and the rotation direction of the object. As well as one-dimensional radar target classification, two approaches can be implemented that are based on raw data or on pre-processed and feature extracted vectors.

Pre-Processing Techniques. The basic procedure of discrimination, as described by Musman et al. (1996), consists of selecting a threshold to remove the background noise and clutter from the original image by target normalisation, and subsequently smoothing the image. Target normalisation can be seen as the extension of the one-dimensional pre-processing previously described into a two-dimensional environment. To smooth the image, *Neighborhood-averaging* processing is applied to each pixel of the $N \times M$ image (Ning et al., 2003). Considering the target image $f(x, y)$ and the smoothed one represented by intensities $g(x, y)$, where $x = 1, 2, ..., N$ and $y = 1, 2, ..., M$, then the pixel intensity $g(\overline{x}, \overline{y})$ depends on the averaged value of the neighborhood of $(\overline{x}, \overline{y})$ as follows:

$$g(x, y) = \frac{1}{D} \sum_{(i,j) \in S} f(x, y) \tag{2.1}$$

where S is the set of coordinates of those points which are located in the neighborhood of $(\overline{x}, \overline{y})$ except the point $(\overline{x}, \overline{y})$, and D is the number of the elements of S. Image recombination techniques such as Multi-Look processing (see Chapter 1) can be thought of as noise, speckle and multipath reduction algorithms. Other pre-processing techniques such as Spatially Vari-

ant Apodization (SVA, Stankwitz and Kosek (1995)), High Definition Vector Imaging (HDI, Benitz (1997)) and image interpolation (Owirka et al., 1999) can be used in order to increase the resolution and more accurately locate the scattering centres.

Feature Extraction. Feature selection for ISAR images results a difficult but important task and the quality of the successive feature extraction is related to the accuracy and speed performance of the ultimate target classification. Possible features can be *geometrical moments* (e.g. Botha et al. (1996)), *outline extraction* (e.g. Ning et al. (2003) and Vespe et al. (2006a)) or *shape and size features* such as orientation, eccentricity and axes length (e.g. Vespe et al. (2005c)). As for the one-dimensional case, the location and intensity of the main scattering centres may be used as features extracted from images, again using superresolution algorithms or main peaks information (e.g Novak et al. (1996)).

Principal Components Analysis. In a typical pattern recognition problem, it is often necessary to reduce the data dimension of the input of the classifier. This is mainly due to an intrinsic degree of redundancy within the data where there are regions where no information is contained (such as a shadow). It also helps to reduce the very high quantities of data that otherwise have to be processed. Thus we may consider range profiles or two-dimensional images as providers of feature vectors that are to be separated. After focusing on those range resolution cells or pixels whose intensities depict the target backscattering, the number of elements representing the backscatter can be reduced with an information loss which is assumed negligible. Furthermore, dimensional reduction of this type also attempts to emphasise the differences between patterns and hence enhance classification performance.

Principal Components Analysis (PCA), also known as Karhunen-Loève expansion, (Jolliffe, 1986) is a statistical method that enables the data to be represented in a different vector basis such that it is possible to remove similarities (which therefore do not contribute to the classification process). After subtracting the mean $\bar{f}$ from each of the vectors of the training set F and producing a zero-mean set of data, the covariance matrix C can be formed:

$$\mathrm{Cov}(F) = \frac{1}{N} \sum_{n} \left(f - \bar{f}\right) \left(f - \bar{f}\right)^{T} \tag{2.2}$$

After calculating the eigenvectors of $\mathrm{Cov}(F)$, the P most significant in terms of largest

eigenvalues are selected to form a new basis vector $V = (v_1, v_2, ..., v_P)$. The test and template feature vectors can then be transformed as follows:

$$f' = V^T \left(f_n - \bar{f} \right) \tag{2.3}$$

The number of principal components $PC = P$ is chosen as a function of the classification rate achieved. This usually becomes stable once the PCs necessary to fully describe the data have been selected. After testing the classifiers, their mean value correct classification rates are analysed versus the number of principal components representing the feature vectors and the lowest number PC that guarantees a sufficient separation capability in the feature space is selected. This processing is further described in Chapter 4.

For two-dimensional classifiers, the pattern after pre-processing consisting of a $M \times N$ matrix of intensities is reshaped into an MN dimensional vector and analysed using PCA.

2.1.3 Classification Problem

Classification of attributes measured by a sensor is a crucial task for a number of applications such as automatic target recognition from radar signatures, automatic speech recognition, classification of text into several categories, automatic recognition of handwritten characters, automatic recognition of images and so forth. In this section the main definitions in classification/recognition problems are presented as well as theory of three approaches to classification. Subsequently, the non-cooperative target recognition is examined both using one and two dimensional images from complex targets.

Pattern Recognition occurs when an element from a population P is assigned to a known set of objects named subpopulation S (Looney, 1998). If the object is assigned to a subpopulation made by just one element, then the recognition is called *identification*. On the other hand, the process of grouping objects together into sub-populations (classes) according to particular characteristics, is called *classification*. Nonetheless the terms classification and recognition are here used under the same meaning that is to assign a category to an input set of attributes.

The area of *Pattern Recognition* deals with classification and recognition procedures and belongs to the field of *Machine Intelligence*, i.e. machine learning and decision-making.

Classification and recognition are usually categorised into a number of areas. The first

is the historical *statistical pattern recognition* which is characterised by an estimation sub-problem of feature values probability densities. This approach, as well as the *rule-based* and *graph theoretic*, is a sub-area of the *decision-theoretic* one. A more recent method of decision-making is called *structural* (or *syntactic*) and is based on the concept of extracting structures from the pattern and using strings of symbols selected from an alphabet to identify them (see Chapter 7). Such strings are considered sentences in a particular grammar whose rules are called *productions*: if an object's structural attributes can be translated into a sentence in a grammar, then it is associated with the class represented by that grammar, otherwise it belongs to another grammar. Ultimately, a modern approach to perform pattern recognition and classification is the *associative mapping* which is based on non-linear mapping T of the N-dimensional cube into an M-dimensional cube: $T : [a, b]^N \rightarrow [a, b]^M$. That is, the feature vector space is mapped into the classes space. Associative mapping is then divided into *Feed-Forward Artificial Neural Networks (FANNs)*, *self-organising networks* and *hybrid networks* (a more exhaustive treatment may be found in Duda et al. (2001), Theodoridis and Koutroumbas (2006) and Bishop (1995)).

Classification is a process based on making a decision on a pattern that can be a multi-dimensional vector of values extracted from an object. Ideally, a feature vector has zero redundance, that is, all the attributes are statistically independent. In other words, the optimal feature vector dimension is achieved when none of the features can be determined by a function of other features in the set or estimated from them as a result of correlation. In general, a set of sensed attributes is converted to a feature vector of lower dimension that contains all the essential information for recognition and classification purposes. Nevertheless, feature vectors extracted from objects belonging to the same class are usually different. This is due to a number of factors such as the natural variation between objects within the same class, noise (measurement and other random errors), systemic error (bias) in measuring instrument, data-acquisition and pre-processing system. For these reasons, the feature vector – according to some probability distribution – appears to be a random variable X assuming values $X = x$ for each sampled sensed object. Sub-populations are therefore characterised by multi-dimensional probability distributions.

Assuming $f_1(x)$ and $f_2(x)$ the probability density functions for the feature Gaussian random variables $X = X^{(1)}$ and $X = X^{(2)}$, this corresponds to class "1" and class "2" respectively. As can be seen from figure 2.2, a one-dimensional feature vector represents the input pattern. By comparing the statistical parameters from a set of *a priori* measurements of the two different classes, the probability of deciding either for the first or the second class can be computed. The dashed line indicates the value for class separation in the feature space: if a sensed attribute x is greater than the separation value, for instance, it is more likely that it belongs to class "2". In a two dimensional feature space, the object linear separation is represented by a straight line.

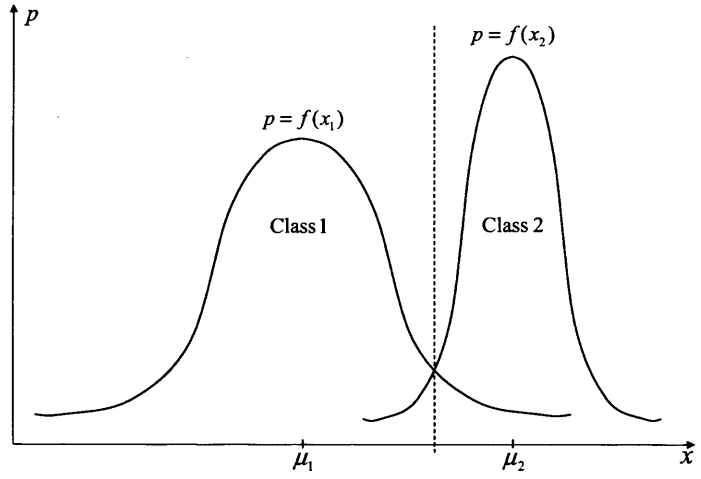

Figure 2.2: *One-dimensional feature space separation of a 2-class problem.*

In figure 2.3 two types of object separation are represented: when it is not possible to separate objects belonging to different classes by a linear subspace, *non-linear* separation occurs. More generally, in a N-dimensional feature space containing two subpopulations, any translation of an $(N-1)$-dimensional linear sub-space by a vector is called a *hyperplane* of N-dimensional space. The separating hyperplane only exists in case of linear separability of objects. If several classes can be separated two at time by hyperplanes they are called *pairwise linearly separable*, also known as *multi-linear separation.*

As described by Looney (1998) *Automatic Pattern Recognition* (APR) is a system that

contains an input sub-system that accepts pattern vectors from feature space, and a decision-maker subsystem that has the function to decide the classes to which the sensed attributes belong.

Three classifiers will be considered in this thesis, while a comprehensive analysis may be found in Tait (2006). These are (i) a Naïve Bayesian classifier, (ii) a K-Nearest Neighbour classifier and (iii) a classifier using Neural Networks (see respectively Appendices A, B, C), covering different areas of Pattern Recognition. The Radial Basis Function Neural Network (RBFNN) classifier is also introduced in Chapter 4. Finally, in Chapter 7, a novel structural approach based on strings is described and tested. These algorithms have been designed to accept multiple perspectives from the target and their performance is widely analysed in Chapters 5, 6, 7, and 8.

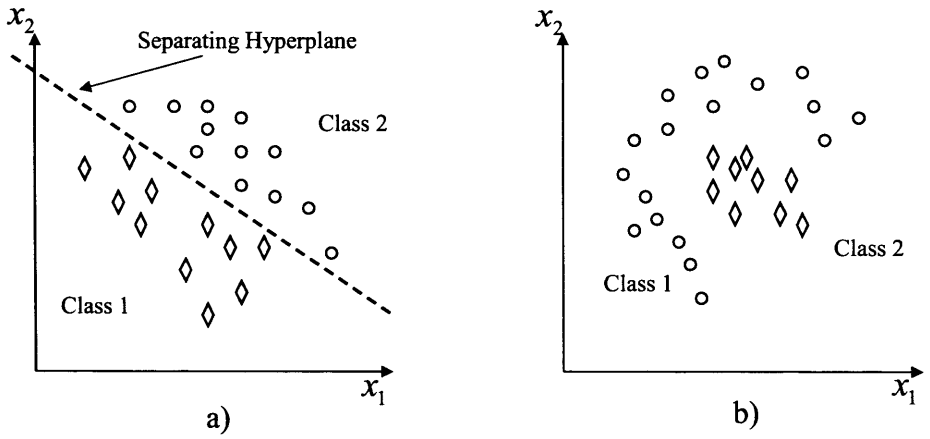

Figure 2.3: *Linear (a) and non-linear (b) two-dimensional feature space separation of two subpopulation.*

The use of an Hidden Markov Models (HMM) multi-aspect classifier has also been implemented (Appendix D) following the methodology developed by Shihao et al. (2005) and compared with the approach investigated in this research work (Chapter 5).

2.2 Summary

In this chapter, the concept of pattern information content has been described for the different stages of pattern recognition. Different techniques for extracting features from one- and two-dimensional radar signatures, although reducing the data redundancy, lead to different target representations and therefore unavoidably affect the classification process. Care needs to be taken when choosing the recognition technique to be used, since different approaches are sensitive to different features. For this reason, during the following chapters, a number of feature extraction and classification approaches are implemented in order to make the analysis of this thesis work focused on the Multi-Perspective benefits on classification performance, independently of the particular technique used.

Chapter 3

Data Processing and Imagery Formation

In this Chapter, the theoretical processing techniques described in Chapter 1 are applied to real turntable data collected from an ISAR system and subsequently preparing the data for feature extraction and recognition.

3.1 Real HRR Data Processing

Data from the *Airborne Data Acquisition System* (ADAS) provided by Thales Sensors has been processed by implementing a frequency domain reconstruction of the target reflectivity spectrum. Firstly, a description of parameters of the system is presented and ultimately, high resolution signatures of calibration scatterers and vehicle target are shown.

The block diagram in figure 3.1 shows a simplified version of the acquisition System. An ISAR geometry has been adopted, where the target is located on a turntable at a fixed and known distance from the system. A waveform generator provides the input signal, a 500 MHz chirp-modulated pulse, which is subsequently mixed with a coherent local oscillator in order to produce the required transmission frequency. The signal is then amplified, linearly polarised (either vertically or horizontally) by the polarisation switch device, and finally transmitted. At the receiver, both horizontal and vertical polarised signatures are passed through a low noise amplifier and subsequently mixed to a lower frequency using the same local oscillator as in the transmitting chain. The system is enabled not only to produce H and V signals but it also distinguishes between cross-polar and co-polar reflection, giving four possible returns from the target and, furthermore, the possibility of fully polarimetric data processing in order to enhance

the eventual characterisation of the signature. Nevertheless, the polarimetric information has not been completely investigated in this research work since the data need to be calibrated.

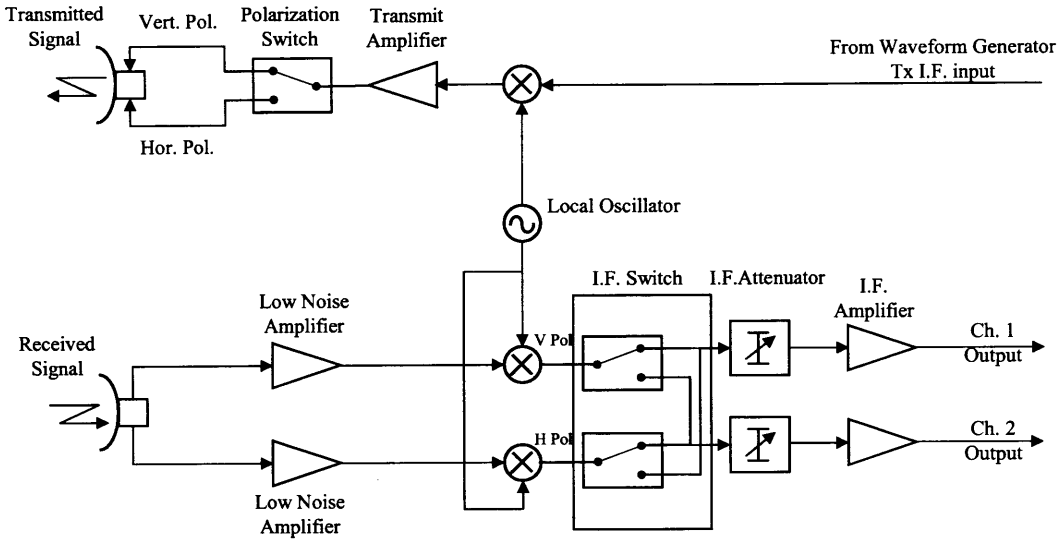

Figure 3.1: *ADAS Block Diagram.*

The Co- and Cross-polarised signatures are obtained by an Intermediate Frequency Switch which can either select Vertical polarisation into channel one (Ch.1) and Horizontal polarisation into channel two (Ch.2) or viceversa. These signals are then passed through an attenuator and amplifier stage in order to adapt the signal strength within the dynamic range of the acquisition system. Signals are then passed to an Analogue to Digital Converter (ADC) and finally stored.

ADAS data are Stepped-Frequency Chirps (see Section 1.1.2). A single chirp-modulated pulse signature (Table 3.1) is now considered. After pulse compression, the range resolution becomes 30 cm, that is $1/PCR$ times the resolution achievable without chirp-modulating the signal (51.15 metres).

Before presenting a first target scenario, a loop test data signal provided with the ADAS set of data is described. It is a recording of the transmitted pulses and it is created by missing out the transmit and receive antennas, that means errors in the antennas will not be considered. In order to evaluate the degree of system calibration and the distortions given by the other

components, the measurement noise at the receiver needs to be evaluated and its probability density function parameters investigated.

Pulse Length	T	341 ns
Pulse Repetition Interval	PRI	2.048 μs
Centre Frequency	f_c	9.25 GHz
Chirp Rate	b	$1.46 \cdot 10^{15}$
Pulse Compression Ratio	PCR	170.5
Transmitted Bandwidth	B_{sw}	500 MHz
Sampling Frequency	f_s	500 MHz
Range Resolution	Δr	30 cm

Table 3.1: *ADAS Single Chirp-Modulated Pulse Parameters.*

Transmitted and received signals are sampled at the sampling rate $f_s = 500$ MHz and consist of 1024 samples (one every 2 nanoseconds).

Time-domain and frequency domain representations of the loop signal are shown in figure 3.2. From the quadratic phase trend and the time-frequency shape of the spectrogram, it can be concluded that the modulation in frequency can be assumed approximately linear. Nevertheless, the frequency domain representation underlines higher transmission power at higher frequencies, provoking an imperfect linear shape in the spectrogram. Furthermore, a folding effect at about 0.5 μs can be seen in the spectrogram shown by the residual frequency components of the chirp modulation. This is not visible for the lower frequencies because of their lower power.

The amplitude and phase representations of the transmitted pulse highlight the residual presence of noise at the receiver. Real and imaginary parts (in-phase I_n and quadrature Q_n components) of noise have been extracted in the time domain from the "silent" time length. Their probability densities have been calculated, showing both in-phase and quadrature components have zero mean and similar variance.

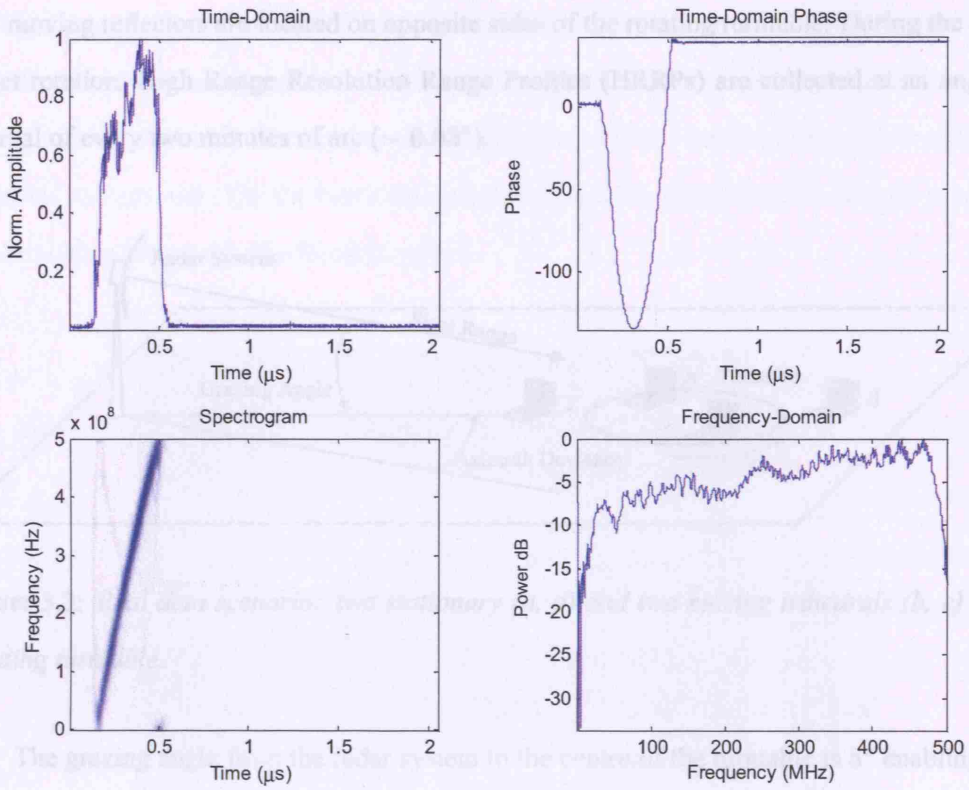

Figure 3.2: *Time and frequency representations of averaged transmitted chirp-modulated pulses.*

As discussed by Papoulis (1984), if both random processes I_n and Q_n are statistically independent and Gaussian distributed, with same variance and zero mean, and provided that the phase is uniformly distributed, then the joint probability density function is Rayleigh distributed. In conclusion, the noise is White-Gaussian, continuous, it exhibits Rayleigh-distributed amplitudes and, therefore, it can be assumed as measurement noise only. As a result, it can be compensated for by averaging a series of recorded loop signals in order to produce the reference for successive matched filter operations.

The first target scenario of real measurements consists of four trihedral corner reflectors (figure 3.3). Two stationary trihedral targets are located in front and behind a turntable and provide a means for computation of the radar point spread function. For collection of the signatures for classification the two rotating corner reflectors are replaced by full scale vehicles.

The moving reflectors are located on opposite sides of the rotating turntable. During the 360° target rotation, High Range Resolution Range Profiles (HRRPs) are collected at an angular interval of every two minutes of arc ($\sim 0.03°$).

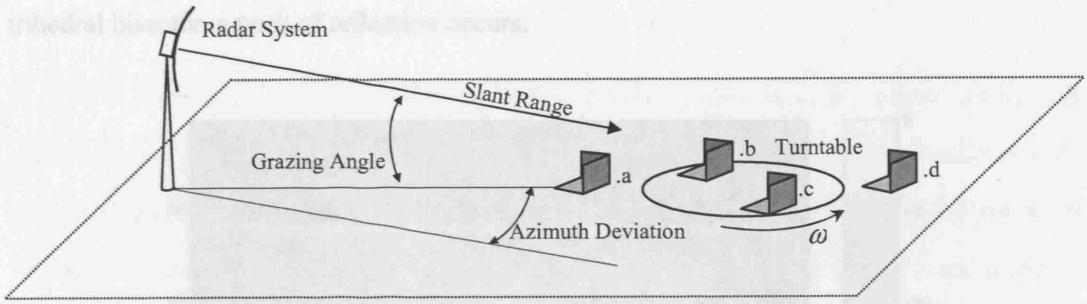

Figure 3.3: *Real data scenario: two stationary (a, d) and two moving trihedrals (b, c) on a rotating turntable.*

The grazing angle from the radar system to the centre of the turntable is 8° enabling the radar to illuminate the target far range scatterers and reducing their occlusion. On the other hand, range measurements are the projection on the line of sight of the actual plane where scatterers are positioned (*slant range*). Finally, the range from antenna to turntable centre is approximately 80 metres.

After the convolution between the transmitted signal and the target reflectivity function, the received signal in phase and quadrature components are collected by the system and processed to extract the high resolution range profile. For a single chirp pulse (as previously described) pulse-compression is obtained by matched filtering the received signal with the transmitted chirp.

In figure 3.4, the real ISAR data depicting the intensity modulated time history of HRR profiles is shown from the two stationary and the two rotating trihedral reflectors for the geometry in figure 3.3. This image, also known as *sinogram*, describes the scatterers range variation as the target rotates. The amplitude of the sinus representing a trajectory is proportional to the scatterer velocity. Therefore, a straight line corresponds to a scatter presenting zero velocity with respect to the sensor (i.e. a point-like target located in the centre of rotation, or a target

having the same vector velocity of the sensor).

The resolution achieved is $\Delta r = c/(2B) = 30$ cm, deteriorated of a factor 1.3 given by the window function. The two stationary trihedrals show a constant response at near and far range as expected. For the two rotating trihedral targets, when the line-of-sight is on the trihedral bisector, a peak of reflection occurs.

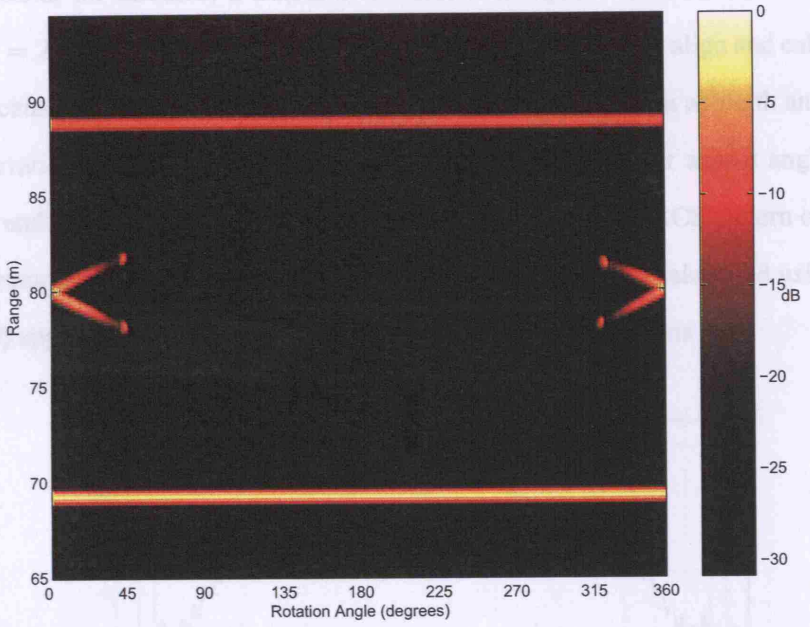

Figure 3.4: *Time history of range profiles from two stationary and two rotating trihedral corner reflectors with* 30 *cm range resolution.*

The reflectivity of a corner reflector at low grazing angles is now investigated. The Radar Cross Section (RCS) σ of a target at a range r from the incident electric field source E_i is defined as the projected area of a metal sphere that would scatter the same power in the same direction as the target (IEEE/ANSI, 1984):

$$\sigma = \lim_{r \to \infty} 4\pi r^2 \frac{|E_s|^2}{|E_i|^2} \tag{3.1}$$

where E_s is the electric field scattered by the target and measured in reception. Trihedral reflectors are calibration targets made up by squared (or triangular) patches orthogonally attached to each other. Trihedral corner reflectors provide a large RCS with an "odd-bounce"

scattering characteristic, usually greater than the actual physical area. For a trihedral corner with square faces l along side, illuminated by an electromagnetic field of wavelength λ, the corresponding maximum RCS can be expressed as follows (Knott et al., 1985):

$$\sigma = \frac{12\pi l^4}{\lambda^2} \tag{3.2}$$

At X band, for instance, a trihedral reflector with square faces 0.5 metres along side, produces $\sigma = 2500\,\mathrm{m}^2 \simeq 34$ dBsm. These reflectors are often used to align and calibrate radar systems because their return is relatively constant over large angles in azimuth and elevation. At zero elevation and for horizontal co-polar returns, the RCS over aspect angle trends of thrihedrals and dihedrals are qualitatively similar. In figure 3.5, the RCS pattern of a dihedral corner reflector is shown with respect to azimuthal angle variation, calculated using Physical Optics (PO) approximation and single and double bounce contributions.

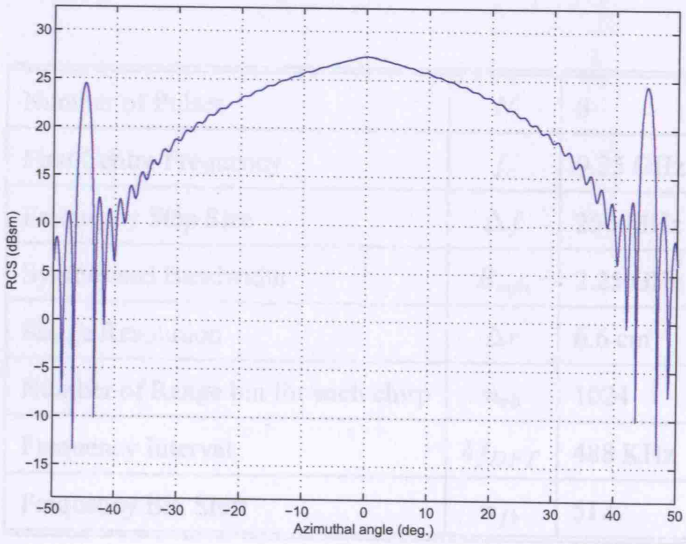

Figure 3.5: *Horizontal co-polar RCS pattern of dihedral reflector having* (50×50) *cm faces. The illuminating frequency is* $f_c = 10$ *GHz and the elevation angle of the incident direction equal to zero.*

The broad central part is due to the double-bounce scattering mechanism, that means the incident wave is reflected once from each face. The peaks at either side of the pattern

are a direct consequence of the specular reflection of one of the faces. For azimuth angles greater than 45°, a corner reflector shows a relatively low RCS. The theoretical corner reflector scattering behaviour described reflects the RCS of the thrihedrals in figure 3.4. As the corners rotate, the monostatic backscattered field decreases progressively until a point is reached where there is a peak of specular reflection. This is given by one of the sides making up the trihedral which is orthogonal to the illuminating radar system (i.e. it faces the radar beam and looks like a flat plate reflector).

At increasing rotation angles the RCS of the target drops since the orientation of the trihedral is such that it tends to reflect incident radiation away from the radar.

The angular dependency of the RCS of a well known reflector such as a trihedral begins to illustrate how the backscattering properties of real targets may vary with the orientation of observation. For example if a target has part of its structure that mimics a trihedral it will only show this feature over a limited angular range.

Number of Pulses	N	8
First Centre Frequency	f_c	9.25 GHz
Frequency Step Size	Δf	250 MHz
Synthesised Bandwidth	B_{syn}	2.25 GHz
Range Resolution	Δr	6.6 cm
Number of Range bin for each chirp	n_{rb}	1024
Frequency Interval	δf_{DFT}	488 KHz
Frequency Bin Shift	n_{fb}	512

Table 3.2: *ADAS stepped-frequency chirp parameters.*

Thus in a Multi-Perspective environment, different angular samples of a targets signature should improve the likelihood of observing a corner or corner-like reflector, which may be an important characterisation for classification. Such shapes can be common on many manmade structures and are often quite dominant features that may prove useful for classification.

We now examine the Stepped-Frequency processing on a burst of eight chirps overlapping for half their bandwidth.

The main specifications from ADAS data format are summarised in Table 3.2. For the chirp specifications of Table 3.1 and equations 1.16 and 1.17, the sub-spectra are shifted in the frequency domain and subsequently combined in accordance to equation 1.15, synthesising a full bandwidth $B_{syn} = 2.25$ GHz, which yields a nominal best range resolution of 6.6 cm. This is demonstrated later on when the resolution achieved is discussed. By following the procedure described in Section 1.1.2 a sequence of high resolution range profiles from the Stepped-Frequency compressed data is shown in figure 3.6.

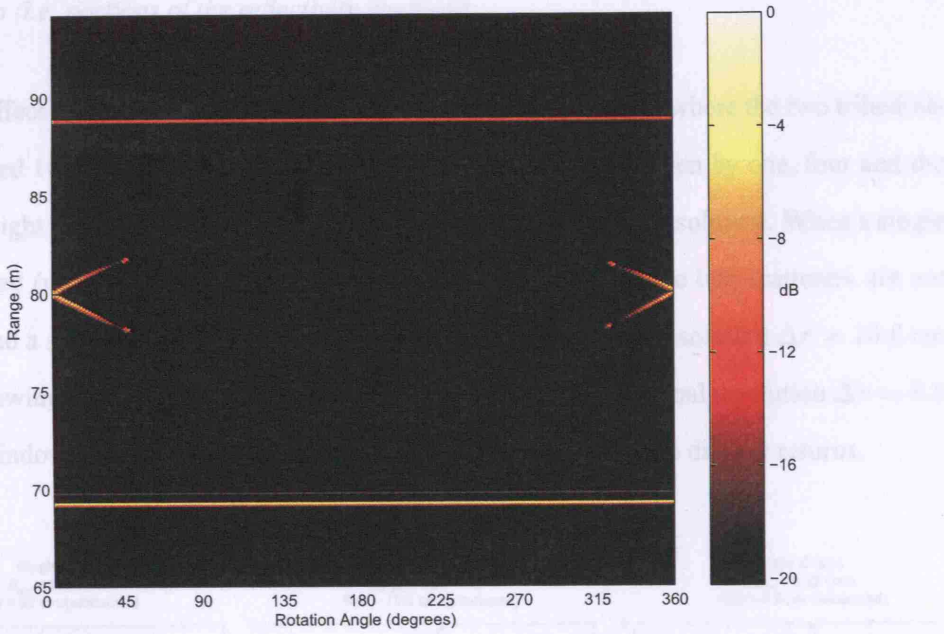

Figure 3.6: *Time history of range profiles from two stationary and two rotating trihedral corner reflectors with* 6.6 *cm range resolution.*

The higher resolution allows now the reduction of the aspect angles where the two corners overlap, and also a more accurate measure of their position. For instance, the specular reflections can now be measured at different aspect angles for the two corners, meaning that the two trihedrals are not perfectly aligned on the turntable when they appear at the same range.

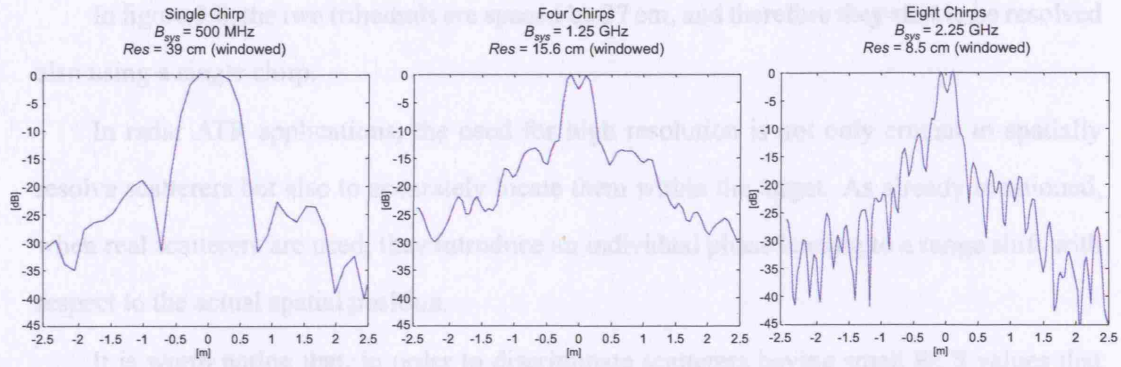

Figure 3.7: *Effects of the bandwidth in resolving scatterers in down range: the thrihedrals are separated by 16 cm while the resolutions synthesised are given using a variable number of sub-spectra (i.e. portions of the reflectivity function).*

The effect on nominal resolution can be derived from figure 3.7, where the two trihedrals are separated 16 cm. The spectrum portions used are respectively given by one, four and the full set of eight Stepped-Frequency chirps, achieving different range resolution. When a single pulse is used (nominal resolution $\Delta r = 39$ cm after windowing), the two scatterers are not resolved and a single mainlobe appears. When four chirps (nominal resolution $\Delta r = 15.6$ cm after windowing) and the full set of compressed pulses are used (nominal resolution $\Delta r = 8.5$ cm after windowing) the two scatterers are progressively resolved into distinct returns.

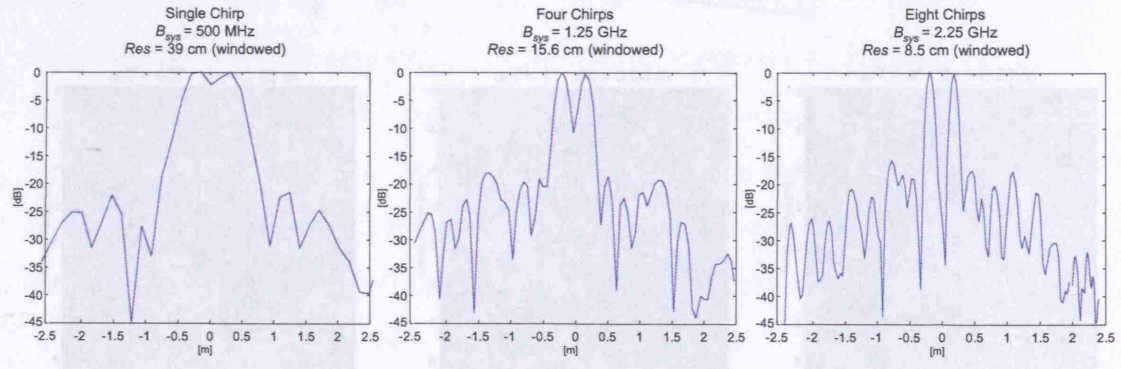

Figure 3.8: *Effects of the bandwidth in resolving scatterers in down range: the thrihedrals are separated by 37 cm while the resolutions are given by a variable number of sub-spectra.*

In figure 3.8, the two trihedrals are spaced by 37 cm, and therefore they start to be resolved also using a single chirp.

In radar ATR applications, the need for high resolution is not only crucial to spatially resolve scatterers but also to accurately locate them within the target. As already mentioned, when real scatterers are used, they introduce an individual phase leading to a range shift with respect to the actual spatial position.

It is worth noting that, in order to discriminate scatterers having small RCS values that are close to others having larger reflectivity, a large dynamic range and low range sidelobes are required. As can be observed from figure 3.7 and 3.8, the disadvantage of Stepped-Frequency techniques is the increase of sidelobes given by the phase discontinuities of the sub-spectra, although partially reducible (e.g. Hai-bin et al. (2005)).

3.1.1 2-D Imaging on Real Data

The 2-D DFT procedure described in Section 1.4 is here applied to a first set of real ADAS data. The scenario is characterised by two stationary and two rotating trihedrals and it has been investigated in the previous chapter.

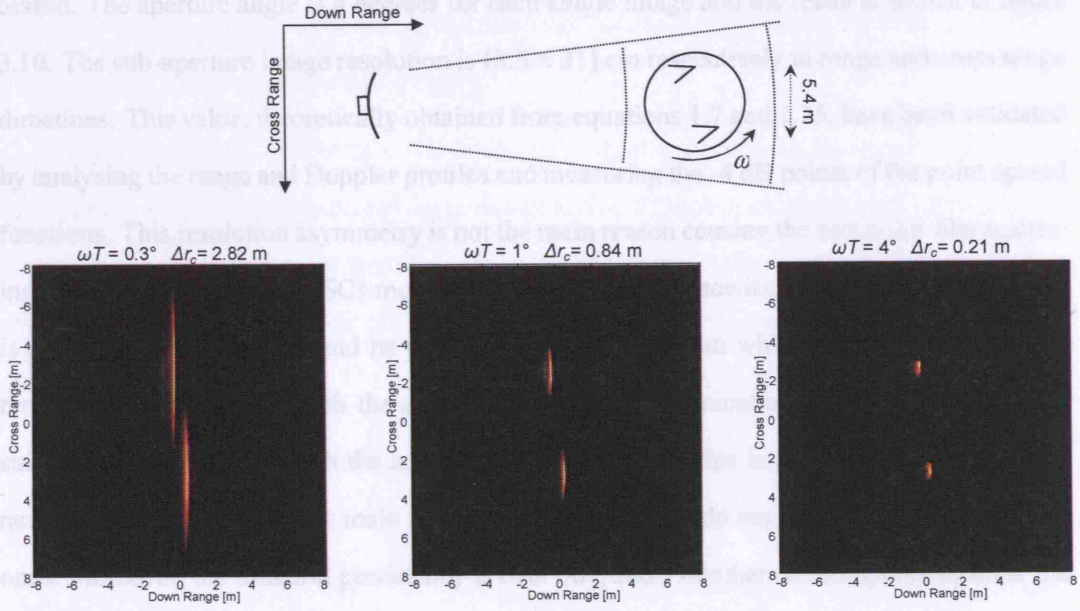

Figure 3.9: *ISAR images from real data at different cross range resolution.*

Each ADAS burst return is collected every $0,034°$. After reconstructing in the frequency domain the sequence of high resolution range profiles, a 2-D DFT has been applied. Figure 3.9 illustrates ISAR images from high resolution range profiles obtained by synthesising three different apertures.

By knowing the step angle between two consecutive received bursts and assuming this angle constant, it is possible to deduce that the synthesised apertures in figure 3.9. The two rotating scatterer phase centers appear in different slant range bins and are equally spaced from the zero Doppler frequency as a consequence of the equal distances from the centre of rotation of the turntable. They become progressively more separated in cross range.

The rotating corners have been isolated from the Zero-Doppler Clutter (ZDC) return, i.e. the stationary calibration trihedrals have been also removed from the ISAR image. ISAR turntable data are usually affected by stationary clutter. For this reason ZDC estimation and subtraction has been performed in the frequency domain (see Chapter 2).

3.2 Radar Image Reconstruction on Real Data

By combining $K = 90$ images from the trihedrals data, a 360 degrees image has been processed. The aperture angle is 4 degrees for each single image and the result is shown in figure 3.10. The sub-aperture image resolution is (8.5×21) cm respectively in range and cross range directions. This value, theoretically obtained from equations 1.7 and 1.25, have been validated by analysing the range and Doppler profiles and measuring the -4 dB points of the point spread functions. This resolution asymmetry is not the main reason causing the non point-like scattering return: according to the SCs model (Rihaczek and Herschowitz, 2000), a corner reflector is a *wave trapping* feature, and its phase center (position from which the feature appears to reradiate) spatially shifts with the aspect angle and the illuminating frequency. This phase shift is ideally still rigid with the actual corner position but this is not true for irregular corners or cavities and this is the main reason why radar images do not usually reflect the optical ones. Moreover, the trihedral persistency is over $90°$, and a number of sub-apertures show the return from outside the main beam since the material is an unperfect absorber, with the result of lowering the quality of the non-coherent superposition.

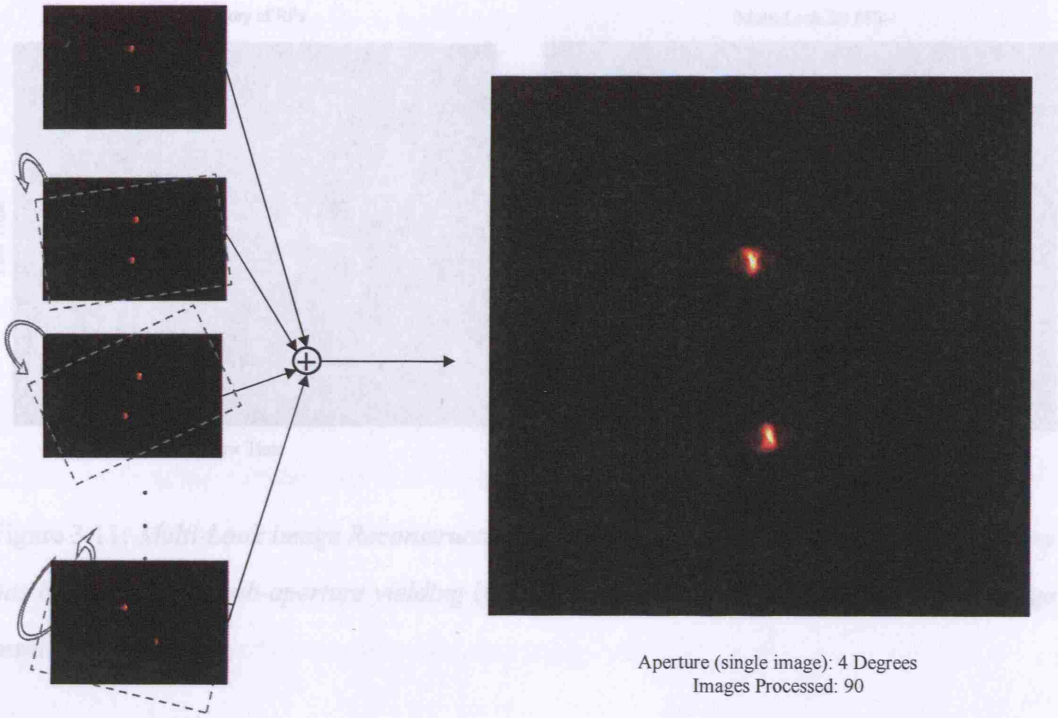

Figure 3.10: *Enhanced Image Processing: Trihedrals scenario.*

With "quality" we mean the extent of resemblance between radar and optical image. Nonetheless it is necessary to clarify that, for classification purposes, to look "more pleasing to the eye" is not the primary requirement.

The information content of a 2-D image resides in the detection, separation and location of the scattering centres forming the complex target. The actual position of trihedrals can be effectively described with higher accuracy and their persistency measured using multiple views of the target, information not visible when the images are combined as in the Multi-Look processing.

The Multi-Look image reconstruction of a ground vehicle target from real turntable experiments is illustrated in figure 3.11. The sub-apertures cover $\sim 210°$ because of the length of the data file available. The shadow effects can be observed from the absence of returns from the far range in the history of HRR profiles at particular perspectives and from the lower intensity of one side of the vehicle in the Multi-Look image.

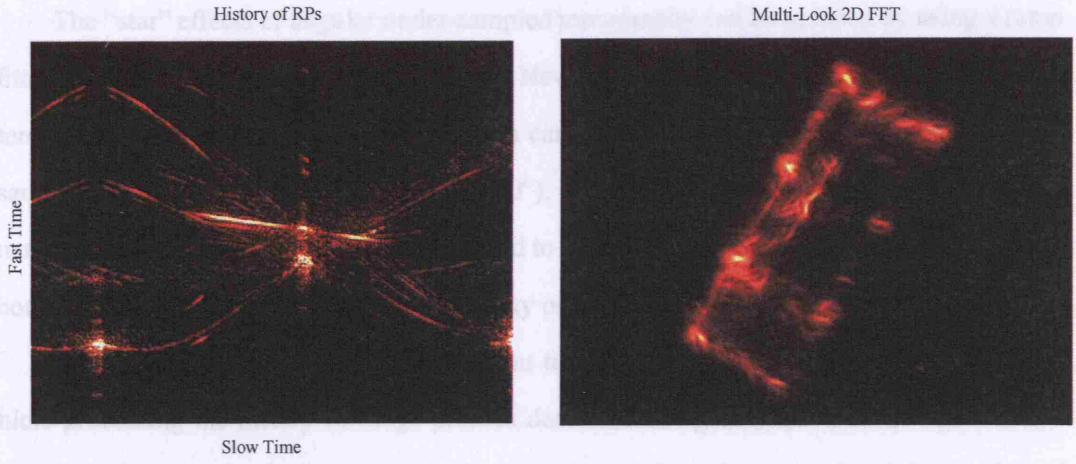

Figure 3.11: *Multi-Look Image Reconstruction from a lorry. The* $\sim 210°$ *HRR profiles history has been divided in sub-aperture yielding* (8.5×21) *cm partial image rage and cross range resolution.*

Figure 3.12 illustrates the back-projection results over 30° of turntable rotation. The position of the corners is the same as depicted in figure 3.9 obtained by 2-D DFT processing.

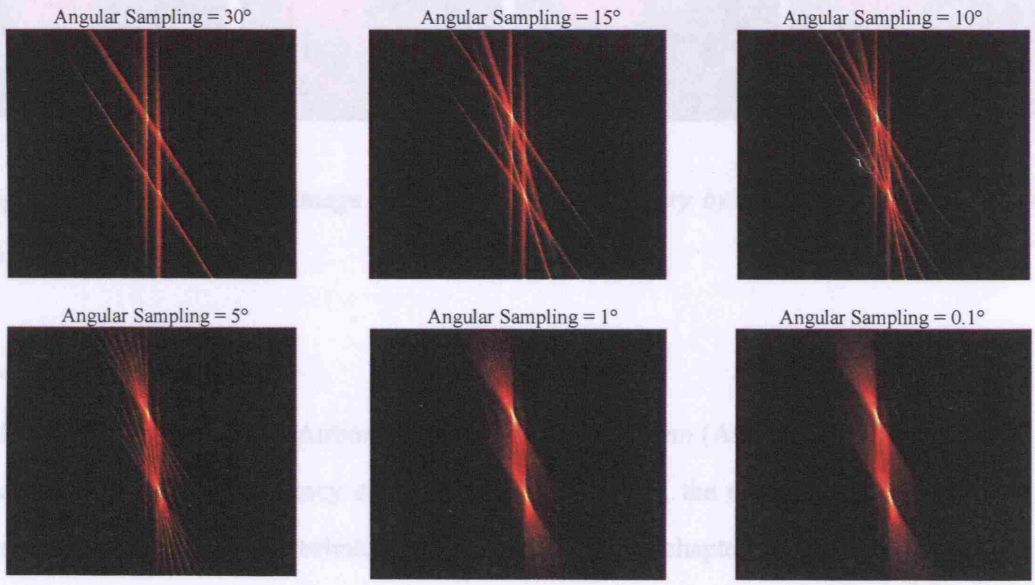

Figure 3.12: *Tomographic radar image reconstruction of an aperture* $\omega T = 30°$ *using different angular sampling steps.*

The "star" effects of angular under-sampled tomography can be reduced by using a ramp filter (*Filtered Back-Projection*, Jain (1989)). Nevertheless, the ambiguities arising from scatterers located at same range from the system can be resolved with a relatively high angular sampling step (one profile every less than 10°). This is significantly dependent on the scenario. Tomographic reconstruction is expected to be subject to ambiguities consistent with the complexity of the target (i.e. the spatial density of the scattering centres) considered.

Figure 3.13 shows different non-coherent tomographic reconstructions of a ground vehicle processing the history of range profiles described in figure 3.11. The ambiguities can be resolved by using a decimation ratio (between the number of processed and the number of available profiles) less than 100, which means an angular step between two profiles of about 3°. Such an angular step would not allow the recovery of the phase information of consecutive range profiles, and therefore to process the data using 2-D DFT.

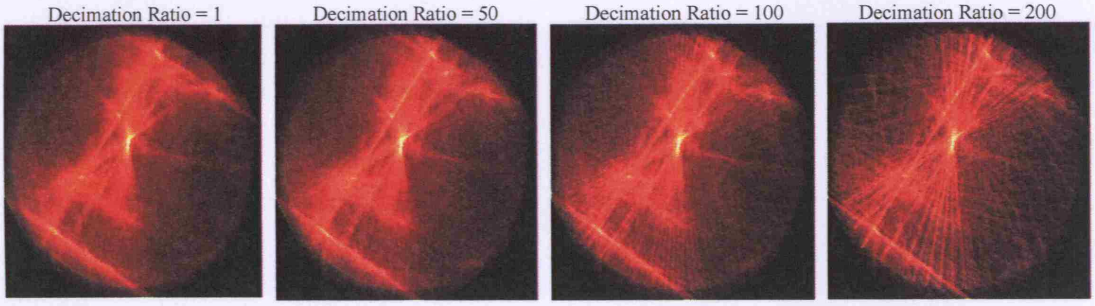

Figure 3.13: *Tomographic image reconstruction from a lorry by using different decimation ratios covering* $\sim 200°$.

3.3 Summary

After the description of the Airborne Data Acquisition System (ADAS) and the radar parameters of the Stepped-Frequency chirp modulation involved, the overview of one- and two-dimensional radar imaging techniques presented in the first chapter have been here discussed and applied to real measurement of ground vehicle targets. The concept of resolution has also been illustrated using real range profiles and images from complex targets. Furthermore, the Multi-Look processing and the Tomographic approach have been introduced as possible ways

to reconstruct a single image from a number of perspectives, which can be represented either by individual range profiles or two-dimensional images. The imagery obtained is the basis of the following chapters on traditional and Multi-Perspective radar target classification.

Chapter 4

ATR Performance Evaluation: Mono-Perspective Results

In this chapter the main parameters are introduced in order to evaluate the classifier performance in terms of probability of correct classification P_{cc}, probability of declaration P_d, false alarm P_{fa}, reliability R and Receiver Operator Characteristic (ROC) curves. These parameters are not always accepted or acknowledged in literature, yielding ambiguous and not exhaustive performance analysis in terms of classification. We intend here to give a definition of the classification requirements and provide for an interpretation of the output of a classifier. Subsequently, after presenting the analysis of one-dimensional and two-dimensional datasets, the results using a single view of the target are evaluated by using different classification techniques and feature extraction algorithms.

4.1 Template Library and Test Set

Two main sets of data are involved in the recognition process: a group of patterns that are used as a reference (*template set*) and a set of patterns used to test the separation capabilities of the classifier (*test set*).

The template set used as a reference by the classifier may be obtained either from radar measurements, CEM simulation or model based features. The selection of the vectors forming the reference library is crucial for the successive classification performance since it will affect the amount of information describing a particular class of targets. When the model or CEM

based approaches are used, the database can be generated using as many templates as needed and the only constraint is the execution time for a decision that increases with the size of the library. Conversely, when a template based methodology is used, the main challenge for the library generation is how to represent the classes when a different system is used: the test set is usually collected by another sensor or by the same one but in different operating conditions (e.g. temperature, weather, clutter distribution and other sources of signature variability) with the result of affecting and possibly deteriorating the classification performance. Ideally, the template database should 'describe' the target classes in a way that the input vector X always finds a correspondence in the respective class database (if any). In other words, the posterior probability $P(C_i|X)$ measured has to be greater than a certain rejection threshold in order to make a declaration for the class C_i. If this does not happen, and the template library is reasonably describing each of the classes, then the object is likely to be unknown to the classifier. It is clear that in case of insufficient classes description in the library, the classifier would show a significant drop of performance in terms of class confusion and declaration rates. On the other hand, in order to reduce *overfitting* effects, it is necessary to allow the classifier to generalise, and this can be done by limiting the information describing the class to the essential amount (restriction absent for CEM and model based approaches where the 'useful' information is *a priori* determined). For these reasons, when the ISAR measurements are used (ADAS and MIDAS data), the template set is generated by a few poses of the target, spaced by a rotation angle showing high decorrelation between the signatures.

A variety of aspect angles are chosen in order to accurately represent each class in the feature space. In view of the fact that a one-dimensional signature is a one-dimensional representation of a three-dimensional scenario, if only one range profile is taken as a feature vector, the size of the target varies significantly with respect to the orientation of the target. As a consequence, objects with the same projected size can be confused. The problem can be also addressed with the same extent to two-dimensional classification procedures and to the scattering centres (SCs) information and their persistency, since most of the main SCs show a large return only for a limited amount of aspect angles (e.g. a corner reflector is visible for 90° of aperture in azimuth, usually significantly reduced by occlusion to a few degrees only).

A possible solution to avoid this classification problem is to consider samples from a number of aspect angles for the training set of each class, covering the greatest number of target orientation. The remaining returns are then used to test the classifier as described in Sections 4.4.3 and 4.4.4.

For the SAR (MSTAR data) classification analysis, two datasets at different grazing angles (15° and 17°) are provided, one for training and the other for testing the classifier. It is clear that scatterer migration as a function of the elevation affects the classifier. Moreover, many scatterers may appear either when the target is sensed from one or the other elevation, yielding a relatively subtle change of the scattering centre spatial distribution. On the other hand, it is also true that, in operational classification scenarios, it seldom happens that the sensed target heading is accurately represented in the database. As a consequence (as previously stated) the generalisation capabilities can be tested using different geometries for the template and test sets as described in Section 4.3.1.

4.2 Classification Parameters and Requirements

For n_c classes $C_1, C_2, ..., C_{n_c}$, the database of templates T is the set of the τ_i signatures representing the i-th class T_i, while the set of test samples A is the union of the n_c subsets of test data A_i formed by α_i vectors:

$$\mathrm{T} = \bigcup_{i=1}^{n_c} T_i \quad \text{where} \quad T_i = \{\mathbf{t}_1^i, \mathbf{t}_2^i, ..., \mathbf{t}_{\tau_i}^i\} \tag{4.1}$$

$$\mathrm{A} = \bigcup_{i=1}^{n_c} A_i \quad \text{where} \quad A_i = \{\mathbf{a}_1^i, \mathbf{a}_2^i, ..., \mathbf{a}_{\alpha_i}^i\} \tag{4.2}$$

where $\mathbf{t}_j^i$ and $\mathbf{a}_j^i$ are the j-th vectors from the i-th class in the template and test sets respectively. In order to evaluate the classification performance, the test set is given to the classifier and the resulting outputs are organised in a matrix often referred to as (*confusion matrix*) as shown in table 4.1.

The test set A from the classes $C_1, C_2, ..., C_{n_c}$ is the input of the classifier. The term $b_{i,j}$ indicates the number of inputs from class C_i mapped into class C_j. Therefore, when $i = j$, the element $b_{i,i}$ represents the amount of test vectors correctly classified, while the remaining elements of the row i are a measure of the misclassification rates for the classes $i \neq j$.

Table 4.1: *Confusion matrix.*

		Output			
Input	(α_i)	C_1	C_2	...	C_{n_c}
C_1	(α_1)	$b_{1,1}$	$b_{1,2}$	...	b_{1,n_c}
C_2	(α_2)	$b_{2,1}$	$b_{2,2}$	...	b_{2,n_c}
$\vdots$					
C_{n_c}	(α_{n_c})	$b_{n_c,1}$	$b_{n_c,2}$	...	b_{n_c,n_c}

Using the formalism of the Bayes' Theorem discussed in Appendix A, we can derive the corresponding posterior probability $P(Out = C_j|In = C_i)$ that, having a pattern of class i as input, a declaration is made for class j:

$$P(Out = C_j|In = C_i) = \frac{b_{i,j}}{\alpha_i} \tag{4.3}$$

For simplification we adopt the notation $P(Out = C_j|In = C_i) = P(Out = j|In = i)$. Thus, the probability of correct classification for a class i is $P_{cc}(i) = P(Out = i|In = i)$, while the corresponding probability of misclassification $P_{m-cl}(i) = \sum_{k \neq i} P(Out = k|In = i)$. The average of these probabilities over the n_c classes gives a first description of the classifier performance:

$$P_{cc} = \frac{1}{n_c} \sum_{k=1}^{n_c} P_{cc}(k) \tag{4.4}$$

$$P_{m-cl} = \frac{1}{n_c} \sum_{k=1}^{n_c} P_{mis}(k) = 1 - P_{cc} \tag{4.5}$$

The likelihood $P(In = i|Out = j)$ of making a decision for class j when a pattern of class i is presented as input to the classifier is:

$$P(In = i|Out = j) = \frac{b_{i,j}}{\sum_{k=1}^{n_c} b_{k,j}} \tag{4.6}$$

Eventually, the prior probabilities $P(Out = j)$ of declaring a class j and $P(In = i)$ of having an input from class i can be written as:

$$P(Out = j) = \frac{\sum_{k=1}^{n_c} b_{k,j}}{\sum_{h=1}^{n_c} \alpha_h} \tag{4.7}$$

$$P(In = i) = \frac{\alpha_i}{\sum_{k=1}^{n_c} \alpha_k} \tag{4.8}$$

As introduced by Cohen the *reliability* $R(i,j)$ of a classifier as the parameter related to the degree of trustworthiness when a declaration is made for a particular class j having an input i, it can be expressed as:

$$R(i,j) = \frac{P(Out = C_j | In = C_i)}{\sum_k P(Out = C_j | In = C_k)} = \frac{b_{i,j}}{\alpha_i \sum_k b_{k,j}/\alpha_k} \tag{4.9}$$

Therefore, the reliability of declaring a particular class j is $R(j) = R(j,j)$, and it gives a measure of the degree of confidence that, when class j is declared, the input was correctly classified. For instance, a classifier always making a declaration for the same class j will present $P_{cc}(j) = 1$, but a corresponding reliability $R(j) = 1/n_c$ and therefore a poor trustworthiness of the decision. The reliability of a classifier R_{avg} is the overall degree of the classifier trustworthiness:

$$R_{avg} = \frac{1}{n_c} \sum_{k=1}^{n_c} R(k) \tag{4.10}$$

The classification process typically requires a high probability of correct classification P_{cc} and reliability R_{avg} . This is calculated in terms of Correct Classification Rate (CCR), which is the P_{cc} on a finite number of measurements. The classifier capability to assign any class label to a particular input from class C_i is the probability of declaration $P_d(i)$:

$$P_d(i) = \sum_{k=1}^{n_c} P(Out = k | In = i) \tag{4.11}$$

The overall probability of declaration $P_d = 1/n_c \sum_i^{n_c} P_d(i)$ is therefore the probability that, for any input target having representation in the template library, the classifier makes a decision for one of the known targets. Therefore, the P_{cc} becomes the probability of correct classification given that a declaration has been made. The declaration properties of a classifier are strictly connected to the unknown threshold specified which determines the rejection of unknown targets. By setting different uncertainty thresholds, the corresponding P_d is obtained and finally the relative P_{cc} obtained. As shown in figure 4.1, the first stage of the classifier is to provide for the classes score, a metric to measure the degree of confidence of the posterior probability estimation. The score is based in the template library information, therefore it

highly depends on how the class is represented by the classifier. Subsequently, after fixing a rejection threshold, whenever the score exceeds this value a declaration is made. Otherwise, the target is considered not having the corresponding class in the template library, and the output is labelled as "unknown". The choice of the threshold is described in the next sections according to the classification approach used.

In real world applications, a measure of the degree of confidence of the classifier is often required in terms of declaration. When an "unknown" object is presented to the classifier, the distance to the templates in the feature space should be greater than a certain threshold.

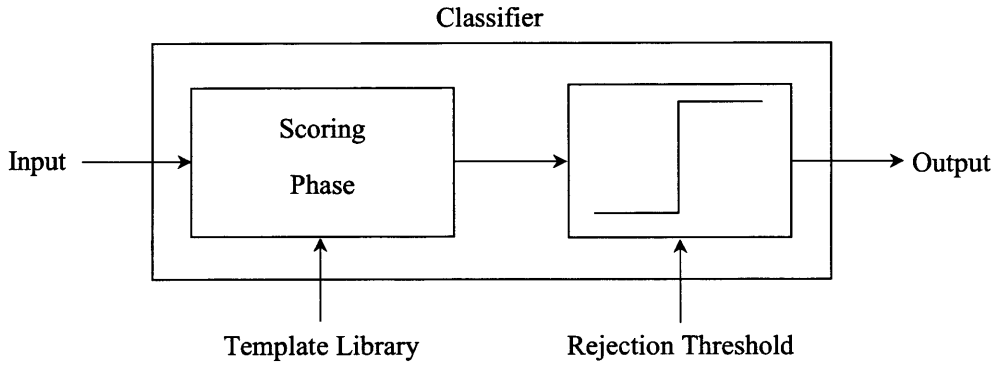

Figure 4.1: *Classification process using a rejection threshold for the "unknown" targets.*

Moreover, in case of similar distance between the input vector and more than one template set T_i, the target should not be declared. In this research work we address the problem of "unknown" targets by denoting a declaration as the decision for one of the available classes. In other words, if the input target has not enough similarity with any target (similarity set by the rejection threshold), then it is not declared. Should the classifier not declare a known input class, then the target would be missed. By defining the probability of missing a target $P_{miss}(i)$ as the number of missed declarations for the test target inputs from class C_i and therefore with a representation in the template database, then:

$$P_{miss}(i) = 1 - \sum_{k=1}^{n_c} P(Out = k|In = i) = 1 - P_d(i) \tag{4.12}$$

The classifier overall probability of missing targets having a represented set of templates

can be expressed as follows:

$$P_{miss} = \frac{1}{n_c} \sum_{i=1}^{n_c} \left(1 - \sum_{k=1}^{n_c} P(Out = k | In = i)\right) = 1 - P_d \tag{4.13}$$

The probability of false alarm P_{fa} can be defined as the probability of making a declaration when an "unknown" target is presented as input to the classifier. Therefore, if the classifier is forced to declare ($P_d = 1$), the uncertainty threshold is set to zero, the probability of missing targets is minimum ($P_{miss} = 0$) although the false alarm rate is maximum ($P_{fa} = 1$).

In table 4.2 the confusion matrix model that has been used to analyse the classifier performance in the following chapters is described. The input classes are shown in the first row of the matrix, together with the corresponding number of samples (α_s) in order to evaluate the degree of balance of the population problem. The parameters previously described are then calculated and placed as introduced in table 4.1. The resulting confusion matrix may be thought of as a set of sub-matrices, related to the different nature of the input class: test, test variant and unknown class.

Table 4.2: *Example of confusion matrix considering test, test variant and "unknown" targets: the P_{cc}, P_{gen}, P_{fa}, P_d and R_{avg} can be evaluated.*

	Output $^{(Reliability)}$				
Input (n_s)	C_1	C_2	...	C_{n_c}	*Unknown*
C_1 (α_1)	$\boldsymbol{Pcc}(1)^{R(1)}$	$P_{m-cl}(1,2)^{R(1,2)}$		$P_{m-cl}(1,n_c)^{R(1,n_c)}$	$P_{miss(1)}$
C_2 (α_2)	$P_{m-cl}(2,1)^{R(2,1)}$	$\boldsymbol{Pcc}(2)^{R(2)}$		$P_{m-cl}(2,n_c)^{R(2,n_c)}$	$P_{miss(2)}$
$\vdots$			$\ddots$		
C_{n_c} (α_{n_c})	$P_{m-cl}(n_c,1)^{R(n_c,1)}$	$P_{m-cl}(n_c,2)^{R(n_c,2)}$		$\boldsymbol{Pcc}(n_c)^{R(n_c)}$	$P_{miss(n_c)}$
C'_1 (α'_1)	$\boldsymbol{Pgen}(1)$	$P'_{m-cl}(1,2)$		$P'_{m-cl}(1,n_c)$	$P'_{miss(1)}$
C'_2 (α'_2)	$P'_{m-cl}(2,1)$	$\boldsymbol{Pgen}(2)$		$P'_{m-cl}(2,n_c)$	$P'_{miss(2)}$
$\vdots$			$\ddots$		
$C'_{n_c^i}$ ($\alpha'_{n_c^i}$)	$P'_{m-cl}(n_c^i,1)$	$P'_{m-cl}(n_c^i,2)$		$\boldsymbol{Pgen}(n_c^i)$	$P'_{miss(n_c^i)}$
U_1 (α''_1)	$P_{fa}(1,1)$	$P_{fa}(1,2)$		$P_{fa}(1,n_c)$	$1 - \sum P_{fa}(1,i)$
U_2 (α''_2)	$P_{fa}(2,1)$	$P_{fa}(2,2)$		$P_{fa}(2,n_c)$	$1 - \sum P_{fa}(2,i)$
$\vdots$			$\ddots$		
$U_{n_c^u}$ ($\alpha''_{n_c^u}$)	$P_{fa}(n_c^u,1)$	$P_{fa}(n_c^u,2)$	...	$P_{fa}(n_c^u,n_c)$	$1 - \sum P_{fa}(n_c^u,i)$

Test target matrix The targets from classes $C_1, ..., C_{n_c}$ are the input to the classifier. They have a representation in the template library and give a measure of the declaration capabilities of the classifier. From this sub-matrix it is possible to measure the correct classification probability (equation 4.4), the probability of declaration (4.11) and the reliability (4.9).

Furthermore, the misclassification probabilities $P_{m-cl}(j,i) = P(Out = j|In = i)$ often resemble similarities between the features from class C_i and C_j, although the opposite is not always true, leading to unsymmetrical test data matrix. This can be explained by the fact that the templates and input vectors from the two classes are not always matching the same characteristics (e.g. orientation and range profile variability sources).

Test variant targets matrix The inputs are from different configurations of the template class (C_i'). They show different features given by some variations like the varying equipments fits of an aircraft (e.g. weapon load, wing geometry), or a tank (e.g. different turret orientations, load or skirts). The ideal classifier would recognise those targets and make a declaration for the corresponding class (C_i). Therefore, the correct classification of those inputs can be thought of as a measure of the generalisation performance achieved $P_{gen}(i) = P(Out = C_i|In = C_i')$. The overall generalisation performance can be expressed as follows:

$$P_{gen} = \frac{1}{n_c^i} \sum_{i}^{n_c^i} P_{gen}(i) \tag{4.14}$$

where n_c^i is the number of test variant classes available. Moreover, the reliability and declaration capabilities can be extracted from this matrix, although those parameters would be affected by the extent of the different features in the target variant.

"Unknown" targets matrix The input targets have no representation in the template set. As a consequence, when the classifier declares them to belong to a particular class C_i, a false alarm occurs. Ideally, the confusing (unknown) targets should not be declared,

decreasing the false alarm probability P_{fa}:

$$P_{fa} = \frac{1}{n_c^u} \sum_{k}^{n_c^u} \sum_{h}^{n_c} P_{fa}(k,h) \tag{4.15}$$

where n_c^u is the number of confusing targets available.

4.2.1 Receiving Operator Characteristic (ROC) Curves

The analysis of the declaration, correct classification and false alarm probabilities is often referred to as the Receiving Operator Characteristics (ROC) curves (Egan, 1975). The ROC curves are a set of operating points expressing the tradeoff between *True Positive* (TP) rate – the ratio of positive instances that are correctly reported as positive – and *False Positive* (FP) rate – the ratio of negative instances that were erroneously reported as positive. In Signal Detection and Machine Learning, the problem is represented by the relationship between 'hit' or detection (TP) and false alarm (FP) rates (Peterson et al., 1954; Jalihal and Nolte, 1994). Finally, the concept can be derived for Radar Target Classification, where the correct classification (TP) is in relation with the false alarm rate (FP) (e.g. Woods and Bowyer, 1997; Mishra and Mulgrew, 2006; Bhanu and Jones, 2000).

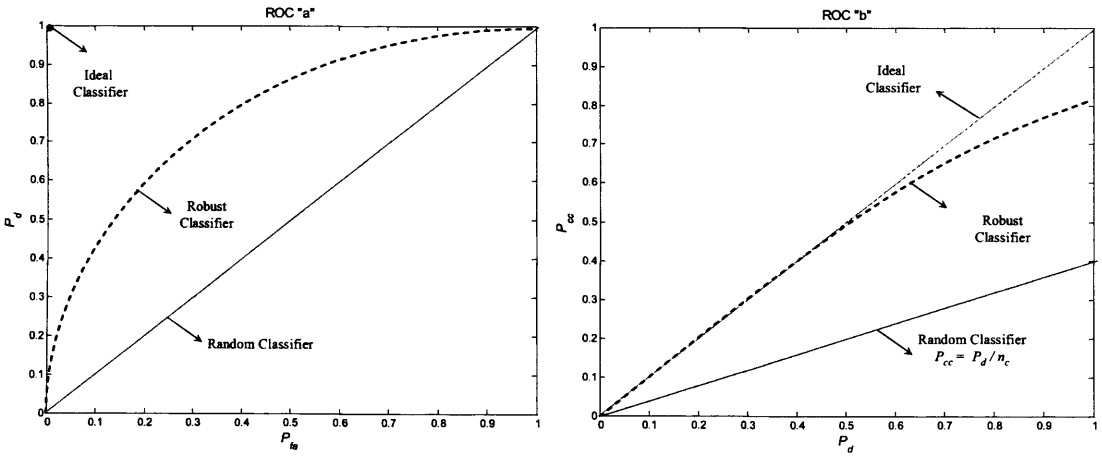

Figure 4.2: *Receiver Operator Characteristic Curves.*

Nevertheless, for ATR problems, the influence of the declaration capabilities is crucial and the analysis can be done by investigating the two ROC curves: $P_{fa} - P_d$ (curve "a") and

$P_d - P_{cc}$ (curve "b") (see Vespe et al., 2006e,g). As previously discussed, a false alarm occurs when a confusing target is declared while only a portion of what has been declared will be correctly classified. As a consequence, a random classifier would present a linear proportion between false alarm and declaration rates, whereas for a robust classifier, the probability of declaration is greater than the false alarm rate.

The ideal classifier would hypothetically be represented by a single point (figure 4.2.a), indicating $P_{fa} = 0$ and $P_d = 1$: every test target is declared and all the confusing input targets are not recognised as belonging to any of the represented classes. Conversely, a random classifier (i.e. randomly labelling the input as either one of the n_c classes or a miss) would linearly increase the false alarm and declaration rates as the rejection threshold decreases. The ROC curve "b" gives a measure of the percentage of declaration effectively correct; since the probability of correct classification upper bound is the probability of declaration, the bisector represents the ideal classifier, where any declaration is correctly made for the relative class. A random classifier would show a linear proportion between declaration and correct classification rate since, provided that a declaration is made, the probability of correct classification is equal to $1/n_c$.

In conclusion, a robust classifier that has separation capabilities is represented by ROC curves in between the ideal and the random case. As a consequence, the Area Under the Curve (AUC) can be thought of as a metric for evaluating the classifier performance at different operating parameters, although it needs to be always related to the ROC curves generating it (Jin and Ling, 2005). This is due to the possibility of ROC curve concavity drops that represent the classifier performance sub-optimality (e.g. in a ROC curve "a", a particular operating P_{fa} shows a drop of declaration capabilities if compared to a lower P_{fa}).

4.3 Classification Results using 2-D MSTAR Data

4.3.1 Data Overview

The population of targets considered here consists of ten military ground vehicles from the X-Band DARPA Moving and Stationary Target Recognition (MSTAR) SAR database (Air Force Research Laboratory website; MSTAR Program Technology Review, 1996). The images are

formed in a spotlight mode covering 360°, achieving a resolution of approximately 30 cm in both slant and cross range. A set of images processed from 17° of depression is used to train the classifier (i.e. to build the template library), while the testing set is formed by 15° depression SAR images.

Template Depression	17 °
Test Depression	15 °
Centre Frequency	9.60 GHz
Bandwidth	591 MHz
Cross and Slant Range Window	-35 dB Taylor
Slant Range Resolution	30.47 cm
Cross Range Resolution	30.47 cm

Table 4.3: *MSTAR data parameters.*

In table 4.3 the main MSTAR data parameters are shown. Although the scattering centres migration starts to be pronounced between the two different elevations, classification can be successfully attempted, simulating a real world scenario where usually the templates database does not fully describe all the possible azimuth and elevation deviations relative to the target.

Figure 4.3: *MSTAR targets Overview.*

Six targets are selected to form the set of training classes (T72, BTR70, BMP2, C2S1, T62, and ZIL131). Two test variant targets (T72ind and BMP2ind), having a representation class in the template set but without being used to train the classifier, measure the generalisation performance in case of targets presenting different features from the ones used to represent the class they belong to. Eventually, two unknown targets (BTR60 and D7), without representation in the training set, give a gauge of the false alarm rate achieved since the classifier should not make a declaration for these targets (figure 4.3).In figure 4.4 the corresponding SAR images of the population set are illustrated at random orientation.

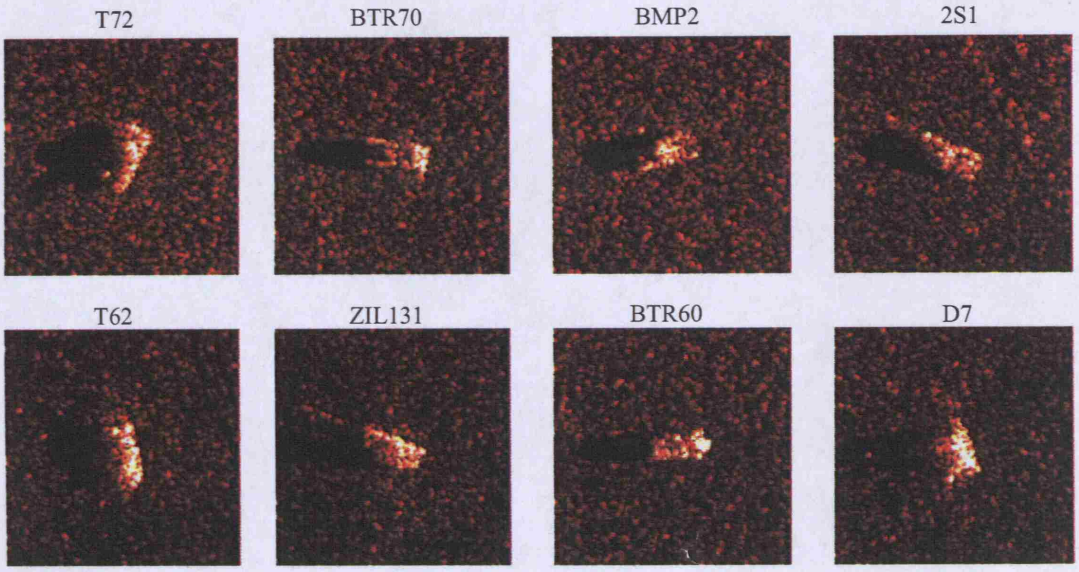

Figure 4.4: *MSTAR SAR images from the test set (15° elevation) at random orientation.*

It is clear that the clutter distribution around each single chip is distinctive and could 'help' the classifier to recognise the target. As a consequence, as described in the next section, the target information has to be extracted and isolated from the clutter.

4.3.2 Preprocessing and Feature extraction

As previously discussed, in a typical pattern recognition problem, it is often necessary to reduce the data dimension of the input of the classifier and increase the separation between targets. The features from the target backscattering and shadow areas are detected in the image

after discriminating the target from its surrounding clutter. The dimension of the classifier input vector is then reduced using Principal Components Analysis (PCA, see 2) on the processed image.

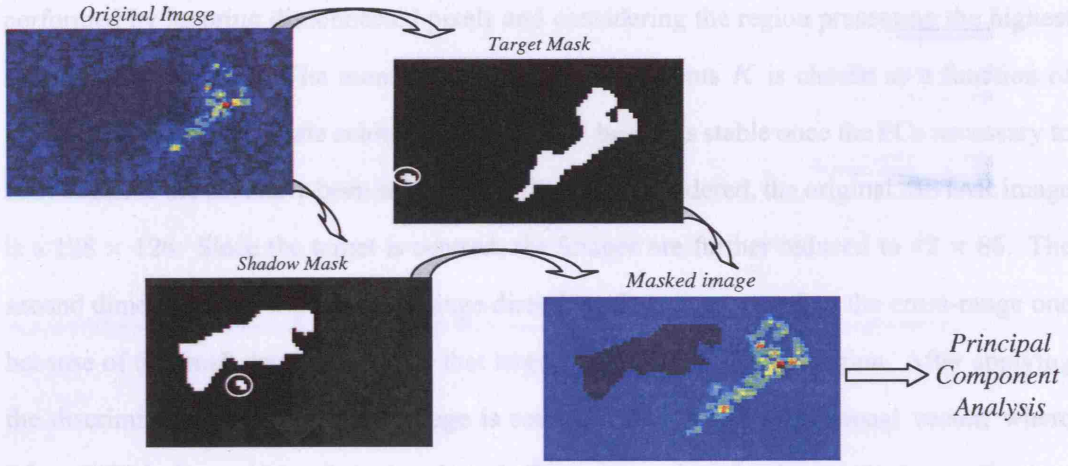

Figure 4.5: *Two different thresholds based on the clutter statistical parameters are applied to the smoothed version of the original image: the target and shadow areas are detected and two masks are formed to isolate the useful information before PCA.*

The need to discriminate the target radar backscattering in SAR images from the clutter return has been exhaustively treated especially in the case of data-sets presenting the same target location in both the training and testing sets, that means also the same clutter content (Schumacher and Schiller, 2005; Novak, 2005). For this reason, the feature vector is formed after clutter cancellation.

Subsequently, the information processed to form the feature vector is extracted from the images after isolating the target area and its shadow using an adaptive procedure that applies two distinct thresholds to the image on the basis of the clutter statistical parameters.

As can be observed from figure 4.5, the result of this discrimination is the detection of the shadow area where the only measurement noise at the receiver is present, and the radar area described by neighborhood of pixels having higher mean and standard deviation if compared with the clutter return. An image mask is eventually obtained and subsequently applied to

filter the original image: the shadow area is represented by constant negative value pixels, the clutter by zero-valued pixels while the target backscattering intensities are preserved as useful information. The intermediate mask images could contain a number of high clutter returns in the target mask as well as their shadow in the other mask (figure 4.5). Segmentation is here performed by ignoring disconnected pixels and considering the region presenting the highest number of pixels only. The number of principal components K is chosen as a function of the correct classification rate achieved. This usually becomes stable once the PCs necessary to fully describe the data have been selected. For the data considered, the original MSTAR image is a 128×128. Since the target is centred, the images are further reduced to 42×85. The second dimension is along the slant range direction. This is greater than the cross-range one because of the small depression angle that largely extends the shadow region. After applying the discrimination algorithm, the image is reshaped into an M dimensional vector, where $M = 3570$ is the number of pixels selected. PCA shows that, for this particular application, 50 principal components are sufficient to stabilise the correct classification rate.

4.3.3 Results

The classification approach chosen uses the Radial Basis Function Neural Network (RBFNN) because of its flexibility and speed of both learning and execution phases. These particular feed-forward neural networks have only three layers (input, hidden, output). Only the neurons belonging to the hidden layer show a nonlinear response, the Radial Basis Function (RBF). In this section, a Gaussian RBF has been used. Each RBF is centered on a small cluster that represents a subclass.

In Table 4.4, the results of the forced decision environment (i.e. the classifier must declare a target class for any input presented), and the correct classification rates of the RBFNN classifier are shown for each target. As can be observed, the classification performance on the independent targets is well below the classification rates of the respective targets used to identify a class. This aspect makes the classification task still unreliable in many applications. As it will be shown in Chapter 5, this effect which leads to upset generalisation capabilities is reduced when another perspective is used to describe the target backscattering.

Table 4.4: *MSTAR data classification results in a forced decision environment* ($P_{fa} = 1$ *and* $P_d = 1$), *yielding* $P_{cc} = 91.6\ \%$, $P_{gen} = 68.5\%$ *and* $R_{avg} = 91.85\ \%$.

		Output $^{(Reliability)}$ (%)						
Input	(n_s)	T72	BTR70	BMP2	C2S1	T62	ZIL131	*Unknown*
T72	(196)	**94.9** $^{(87.7)}$	0.5 $^{(0.5)}$	2.1 $^{(2.0)}$	0.5 $^{(0.6)}$	2.0 $^{(2.1)}$	– $^{(-)}$	–
BTR70	(194)	0.5 $^{(0.5)}$	**98.5** $^{(88.9)}$	– $^{(-)}$	0.5 $^{(0.6)}$	0.5 $^{(0.5)}$	– $^{(-)}$	–
BMP2	(195)	3.6 $^{(3.3)}$	1.5 $^{(1.4)}$	**93.4** $^{(89.2)}$	1.0 $^{(1.2)}$	– $^{(-)}$	0.5 $^{(0.5)}$	–
C2S1	(274)	3.3 $^{(3.0)}$	9.5 $^{(8.6)}$	1.4 $^{(1.3)}$	**80.3** $^{(94.6)}$	5.1 $^{(5.4)}$	0.4 $^{(0.4)}$	–
T62	(273)	5.9 $^{(5.5)}$	0.7 $^{(0.6)}$	4.3 $^{(4.1)}$	1.5 $^{(1.7)}$	**87.2** $^{(92.0)}$	0.4 $^{(0.4)}$	–
ZIL131	(274)	– $^{(-)}$	– $^{(-)}$	3.6 $^{(3.4)}$	1.1 $^{(1.3)}$	– $^{(-)}$	**95.3** $^{(98.7)}$	–
$T72_{ind}$	(195)	**71.3**	2.5	8.7	2.6	14.9	–	–
$BMP2_{ind}$	(196)	14.8	10.2	**64.8**	4.1	5.6	0.5	–
$D7_{un}$	(256)	–	0.4	11.3	20.8	31.0	36.5	–
$BTR60_{un}$	(274)	6.3	49.6	33.2	4.3	1.2	5.4	–

The results for a rejection threshold different from zero and yielding a reduction of the false alarm rate from 100% to 91.5% (-8.5%) are shown in table 4.5.

Table 4.5: *MSTAR classification results at a fixed rejection threshold leading to* $P_{fa} = 91.5\%$, $P_d = 99.05\%$, $P_{cc} = 91.36\%$, $P_{gen} = 67.8\%$ *and* $R_{avg} = 92.47\%$.

		Output $^{(Reliability)}$ (%)						
Input	(n_s)	T72	BTR70	BMP2	C2S1	T62	ZIL131	*Unknown*
T72	(196)	**94.4** $^{(87.7)}$	0.5 $^{(0.5)}$	1.0 $^{(1)}$	0.5 $^{(0.6)}$	2.0 $^{(0.5)}$	– $^{(-)}$	1.6
BTR70	(194)	0.5 $^{(0.5)}$	**98.5** $^{(89.2)}$	– $^{(-)}$	0.5 $^{(0.6)}$	0.5 $^{(0.5)}$	– $^{(-)}$	–
BMP2	(195)	3.6 $^{(3.3)}$	1.5 $^{(1.4)}$	**93.3** $^{(91.8)}$	1.0 $^{(1.2)}$	– $^{(-)}$	– $^{(-)}$	0.6
C2S1	(274)	3.3 $^{(3.1)}$	9.5 $^{(8.6)}$	1.4 $^{(1.4)}$	**79.9** $^{(95)}$	5.1 $^{(5.4)}$	0.4 $^{(0.4)}$	0.4
T62	(273)	5.8 $^{(5.4)}$	0.4 $^{(0.3)}$	3.6 $^{(3.6)}$	1.1 $^{(1.3)}$	**87.2** $^{(91.9)}$	0.4 $^{(0.4)}$	1.5
ZIL131	(274)	– $^{(-)}$	– $^{(-)}$	2.2 $^{(2.2)}$	1.1 $^{(1.3)}$	– $^{(-)}$	**94.9** $^{(99.2)}$	1.8
$T72_{ind}$	(195)	**71.3**	2.5	6.2	1.5	14.9	–	3.6
$BMP2_{ind}$	(196)	14.8	10.2	**64.3**	4.1	5.6	0.5	0.5
$D7_{un}$	(256)	–	–	6.2	17.1	29.6	33.2	**13.9**
$BTR60_{un}$	(274)	5.1	49.2	32.8	3.9	0.8	5.1	**3.1**

The effects are the reduction of P_{cc} (-0.24%), P_d (-0.95%) and P_{gen} (-0.7%) also, although the variation is relatively less significant than the false alarm rate drop. This means that the classifier declares less than for the forced decision implementation but in a more reliable way, since the correct classification variation is lower than the declaration rate one. This can be seen also from the overall reliability increase ($+0.62\%$), demonstrating that the consequence of forcing a declaration shows its effects firstly on those decisions made on the basis of poor confidence. The score produced by the classifier for the confusing input targets is less than the one made for the targets from the test set. Furthermore, when the target is successfully recognised, the classifier score is greater overall than the threshold set, suggesting that the confidence of the decision is stronger than the one of other trials misclassified.

Table 4.6: *MSTAR classification results at a fixed rejection threshold leading to* $P_{fa} = 29.67\%$, $P_d = 81.19\%$, $P_{cc} = 79.61\%$, $P_{gen} = 45.05\%$ *and* $R_{avg} = 98.25\%$.

		Output $^{(Reliability)}$ (%)						
Input	(n_s)	T72	BTR70	BMP2	C2S1	T62	ZIL131	*Unknown*
T72	(196)	**87.8** $^{(94.8)}$	0.5 $^{(0.5)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	11.7
BTR70	(194)	_ $^{(-)}$	**94.8** $^{(96.7)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	5.2
BMP2	(195)	1.5 $^{(1.7)}$	0.5 $^{(0.5)}$	**74.4** $^{(99)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	23.6
C2S1	(274)	0.7 $^{(0.8)}$	2.2 $^{(2.2)}$	0.4 $^{(0.5)}$	**58.8** $^{(100)}$	0.7 $^{(1)}$	_ $^{(-)}$	37.2
T62	(273)	2.6 $^{(2.8)}$	_ $^{(-)}$	0.4 $^{(0.5)}$	_ $^{(-)}$	**72.2** $^{(99)}$	_ $^{(-)}$	24.9
ZIL131	(274)	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	_ $^{(-)}$	**89.8** $^{(100)}$	10.2
$T72_{ind}$	(195)	**52.3**	1	1	1	0.5	0.5	39.5
$BMP2_{ind}$	(196)	8.7	4.1	**37.8**	0.5	1	–	48
$D7_{un}$	(256)	–	–	0	0.7	5.1	11.3	**82.9**
$BTR60_{un}$	(274)	0.8	29.2	10.9	0.4	0.4	0.4	**57.8**

It is also possible to deduce that the classifier is consistently confusing some particular classes: the unknown trials from the class BTR60 is often declared as BTR70 class as it might be expected looking at the geometrical similarities from figure 4.3. The rejection threshold is further increased until achieving $P_{fa} = 29.67\%$ (table 4.6). The correct classification rate, as

a consequence, is progressively reduced, as well as generalisation and declaration rates.

On the contrary, since the classifier is now set to declare for a class only if the degree of confidence is significant with respect to the threshold, the overall reliability increases. This could be an operational scenario where a significant correct classification is preferred.

Table 4.7: *MSTAR classification results at a fixed rejection threshold leading to* $P_{fa} = 3.5\%$, $P_d = 49.1\%$, $P_{cc} = 48.93\%$, $P_{gen} = 16.65\%$ *and* $R_{avg} = 99.7\%$.

		Output $^{(Reliability)}$ (%)						
Input	(n_s)	T72	BTR70	BMP2	C2S1	T62	ZIL131	*Unknown*
T72	(196)	**69.4** (99.3)	_ (–)	_ (–)	_ (–)	_ (–)	_ (–)	30.6
BTR70	(194)	_ (–)	**70.6** (100)	_ (–)	_ (–)	_ (–)	_ (–)	29.4
BMP2	(195)	0.5 (0.7)	_ (–)	**27.2** (100)	_ (–)	_ (–)	_ (–)	72.3
C2S1	(274)	_ (–)	_ (–)	_ (–)	**21.2** (100)	0.3 (1.1)	_ (–)	78.5
T62	(273)	_ (–)	_ (–)	_ (–)	_ (–)	**33** (98.9)	_ (–)	67
ZIL131	(274)	_ (–)	_ (–)	_ (–)	_ (–)	_ (–)	**72.3** (100)	27.7
$T72_{ind}$	(195)	**21.5**	–	0	–	2.6	–	75.9
$BMP2_{ind}$	(196)	2	–	**11.8**	–	–	–	86.2
$D7_{un}$	(256)	–	–	0	–	0.4	–	**99.6**
$BTR60_{un}$	(274)	–	6.6	0	–	–	–	**93.4**

In situations where the need is a low false alarm rate and high reliability for the test targets the classifier declares only when a high score is detected. In table 4.7, the confusion matrix is illustrated for $P_{fa} = 3.5\%$. For instance, the reliability achieved when the class ZIL131 is declared is 100%: only if the input is from the same class, the classifier makes that declaration. The generalisation capabilities are reduced to $P_{gen} = 16.65\%$ but the test variant targets are mostly not declared rather than misclassified. These results can be synthesised using the ROC curves (figure 4.6), according to the model described in Section 4.2.1. The classification performance can be considered robust for both the relationships $P_{fa} - P_d$ and $P_d - P_{cc}$. It is possible to also include in a single graph the generalisation and reliability performance versus the probability of false alarm as shown in figure 4.7: as the rejection threshold increases, the drops of performance in terms of generalisation capabilities, probability of correct classifica-

tion and declaration are counterbalanced by the increase of reliability of the decision.

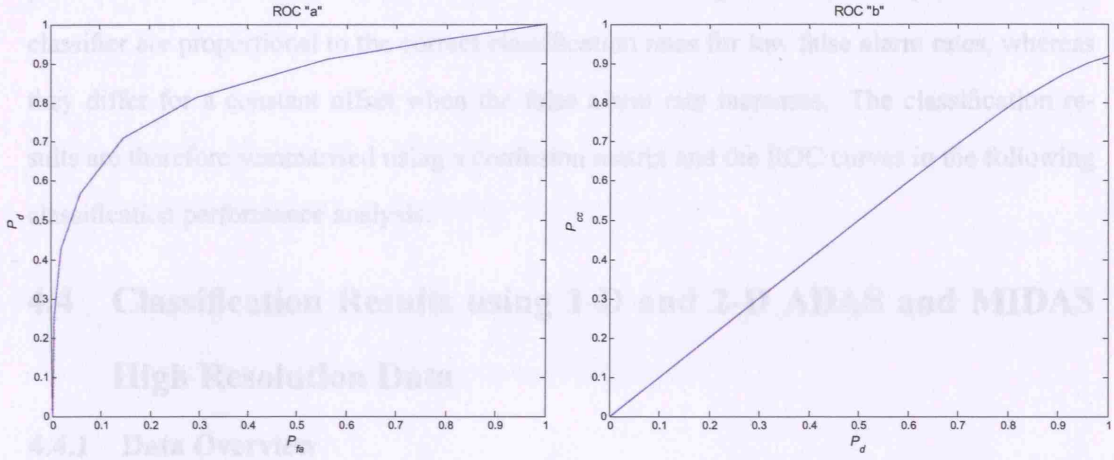

Figure 4.6: *MSTAR single perspective ROC curves.*

Nevertheless, the information contained in figures 4.6 and 4.7 is equivalent: when the reliability is maximum, the difference between declaration and correct classification is minimum, conversely a low reliability would lead to a lower slope of the $P_d - P_{cc}$ graph.

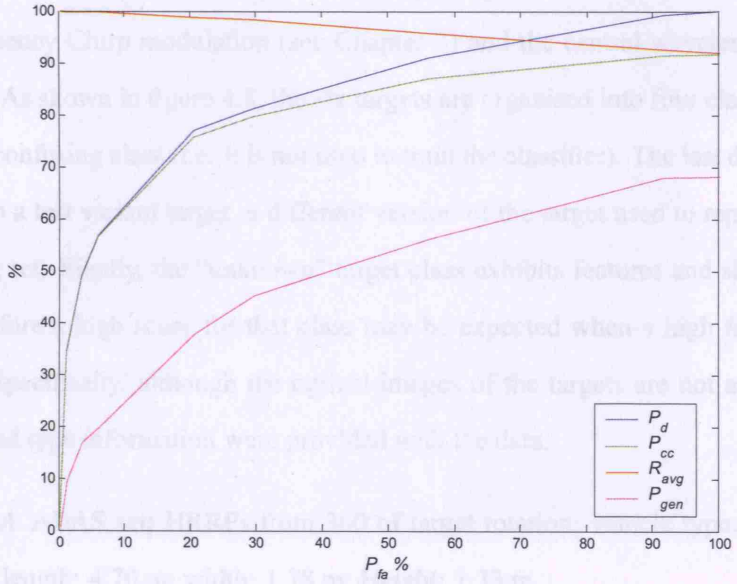

Figure 4.7: *MSTAR classifier performance evaluation.* P_{cc}, P_d, R_{avg} *and* P_{gen} *versus* P_{fa}. *The reliability is proportional to the ratio between* P_d *and* P_{cc}

As a result, the gradient of the line intersecting the origin and the ROC "b" point is proportional to the reliability of the decision. Moreover, the generalisation capabilities of the classifier are proportional to the correct classification rates for low false alarm rates, whereas they differ for a constant offset when the false alarm rate increases. The classification results are therefore summarised using a confusion matrix and the ROC curves in the following classification performance analysis.

4.4 Classification Results using 1-D and 2-D ADAS and MIDAS High Resolution Data

4.4.1 Data Overview

The Airborne Data Acquisition System (ADAS) dataset provided by Thales consists of radar reflectivity frequency samples from three ground vehicles covering 360 degrees of target rotation on a turntable. The reflectivity data from other three targets are obtained by the Mobile Instrumented Data Acquisition System (MIDAS), provided by the ex DERA to "La Sapienza" University.

In both the MIDAS and ADAS systems the high range resolution is achieved by using the Stepped-Frequency Chirp modulation (see Chapter 3) and the central wavelength is about 3 cm (X-band). As shown in figure 4.8, the six targets are organised into four classes (A, B, C, and D) and a confusing class (i.e. it is not used to train the classifier). The last data file (D_{ind}) corresponds to a test variant target, a different version of the target used to represent class D in the training set. Finally, the "unknown" target class exhibits features and shape similar to class A, therefore a high score for that class may be expected when a high false alarm rate is tolerated. Specifically, although the optical images of the targets are not available, some geometrical and type information were provided with the data:

Target Class A ADAS set; HRRPs from 360 of target rotation; vehicle type: Ford Cougar. Chassis length: 4.70 m; width: 1.78 m; Height: 1.33 m.

Target Class B ADAS set; HRRPs from 360 of target rotation; vehicle type: Land Rover; chassis length: 4.44 m; width: 1.68 m; Height: 2.05 m.

Target Class C ADAS set; HRRPs from 360 of target rotation; vehicle type: Tank 55 (turret azimuth $\sim$ 36 degrees, hatches closed, engine off) ; chassis length: 6.20 m; width: 3.60 m; Height: 2.32 m.

Target Class D MIDAS set; HRRPs from 360 of target rotation; vehicle type: Lorry (unspecified, with load); vehicle dimensions not available.

Target Class D_{ind} MIDAS set; HRRPs from 200 of target rotation; vehicle type: Lorry (unspecified, flatbed); vehicle dimensions not available.

Target Class $Unknown$ MIDAS set; HRRPs from 360 of target rotation; vehicle type: Lorry (unspecified, utility passenger car); vehicle dimensions not available.

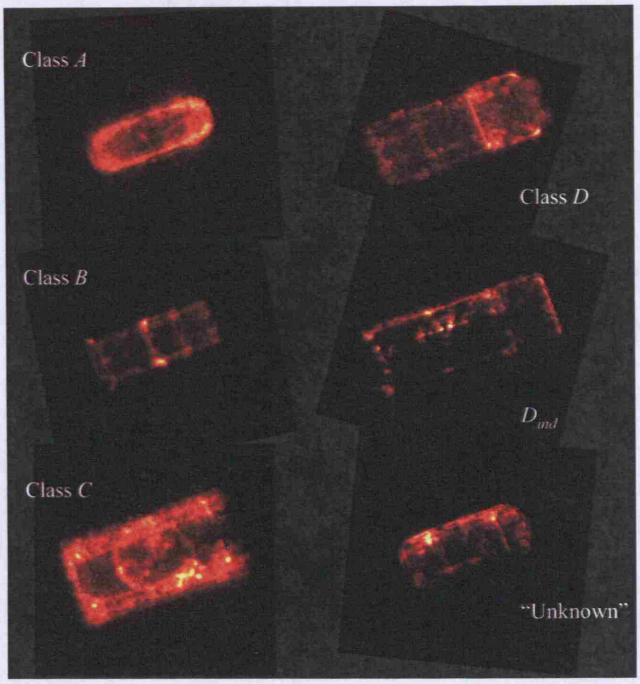

Figure 4.8: *Multi-Look ISAR images over* 360° *(classes A, B, C, D) and* $\sim$ 210° *(D_{ind} and "Unknown").*

4.4.2 Preprocessing and Feature extraction

As explained in Chapter 2, the Zero Doppler Clutter (ZDC) backscattering contribution given by stationary or quasi stationary clutter surrounding the turntable needs to be subtracted from the series of range profiles. Within the ADAS data two different ZDC contributes can be distinguished, the first one common to target A and C, and the second one affecting the profiles from target B. This is because the measurements were collected in two different days, changing the weather conditions (different wind speed and therefore different grass clutter decorrelation velocities) and the clutter distribution (different vehicles manoeuvering around the turntable modifying the gravel).

In figure 4.9.a the series of range profiles having 30 cm slant range resolution are shown versus the turntable rotation for the ADAS targets A, B and C.

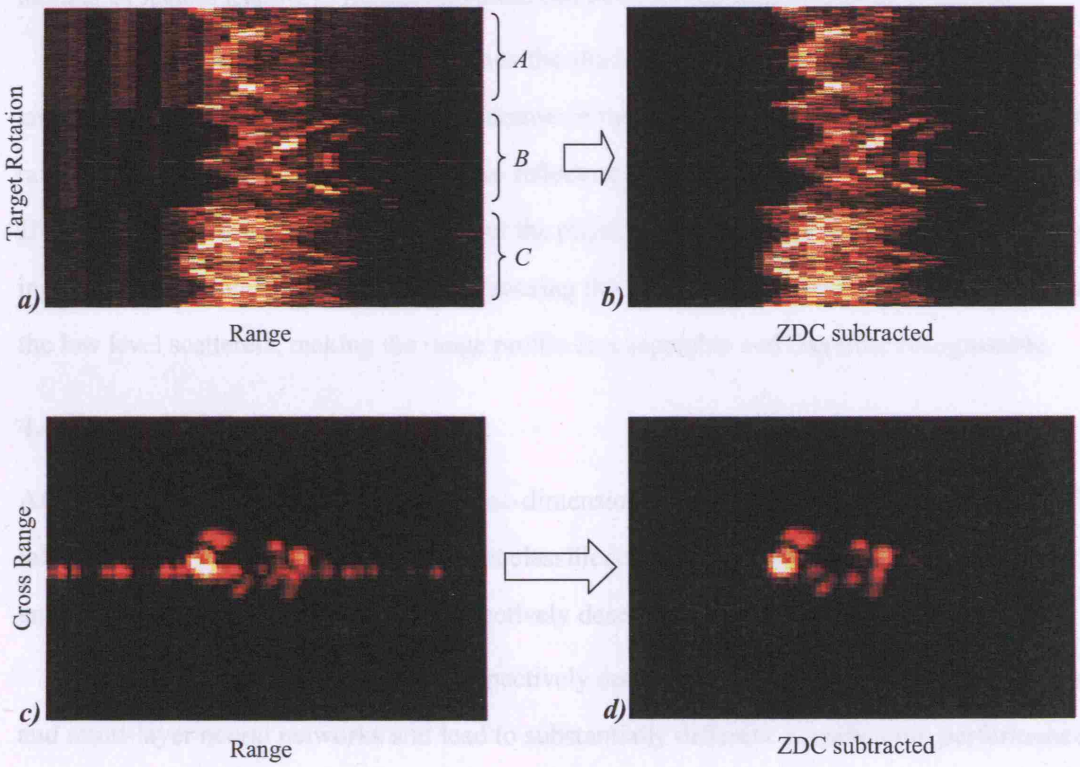

Figure 4.9: *ADAS data ZDC estimation and subtraction results from one-dimensional* (a, b) *and two-dimensional* (c, d) *signatures.*

After the ZDC estimation and subtraction (figure 4.9.b), the target backscattering is clearly isolated. The effect is also appreciable from the two-dimensional ISAR image of the target A in figures 4.9.c and 4.9.d, where a strong intensity line extending throughout the zero-Doppler direction, deteriorating the quality of the region of interest description. Furthermore, since the ADAS and MIDAS measurements were collected in two different locations, they show distinctive and strong ZDC component that need to be subtracted in order to have no influence on the following classification stage. *Non-coherent averaging* and *target normalisation* (see 2) are successively applied in order to increase the SNR.Finally, PCA is applied both on one- and two-dimensional imagery and, after selecting the most significative components, the test set is used to evaluate the classification performance.

Figure 4.10 illustrates the history of HRR profiles of the ADAS-MIDAS targets after Stepped-Frequency compression, achieving $\sim$ 8 cm range resolution due to windowing. A number of factors leading to misclassification can be observed from this representation.

The radar system is at near range, thus the shadowed regions, more pronounced for the targets extended in height (e.g. class C) because of the low grazing angle ($\sim 8°$), appear at far range. Multiple bounces due to corner-like reflectors located at rear of the Land Rover (class B) extend the radar length of the target over the physical length for few degrees of target heading. Besides, for other particular aspects sensing the targets, the specular reflections obscure the low level scatterers, making the range profile less separable and therefore recognisable.

4.4.3 One-Dimensional Results

After PCA processing, the results using one-dimensional images from the ADAS and MIDAS targets are now shown using three different classifiers: *Naïve Bayesian, K-Nearest Neighbour*, and *Feed-forward Neural Networks*, respectively described in appendices A, B, and C.

The different approaches involve respectively decision theory, non parametric techniques and multi-layer neural networks and lead to substantially different classification performance as eventually shown in the ROC curves comparison at the end of this subsection. Nevertheless, as explained in Chapter 5, their Multi-Perspective implementation is similarly effective in suggesting the contribution of the number of perspectives to correct classification indepen-

dently of the pattern recognition approach followed. The high range resolution ($\sim$ 8 cm) range profile history for each target is divided into a template library and a test set.

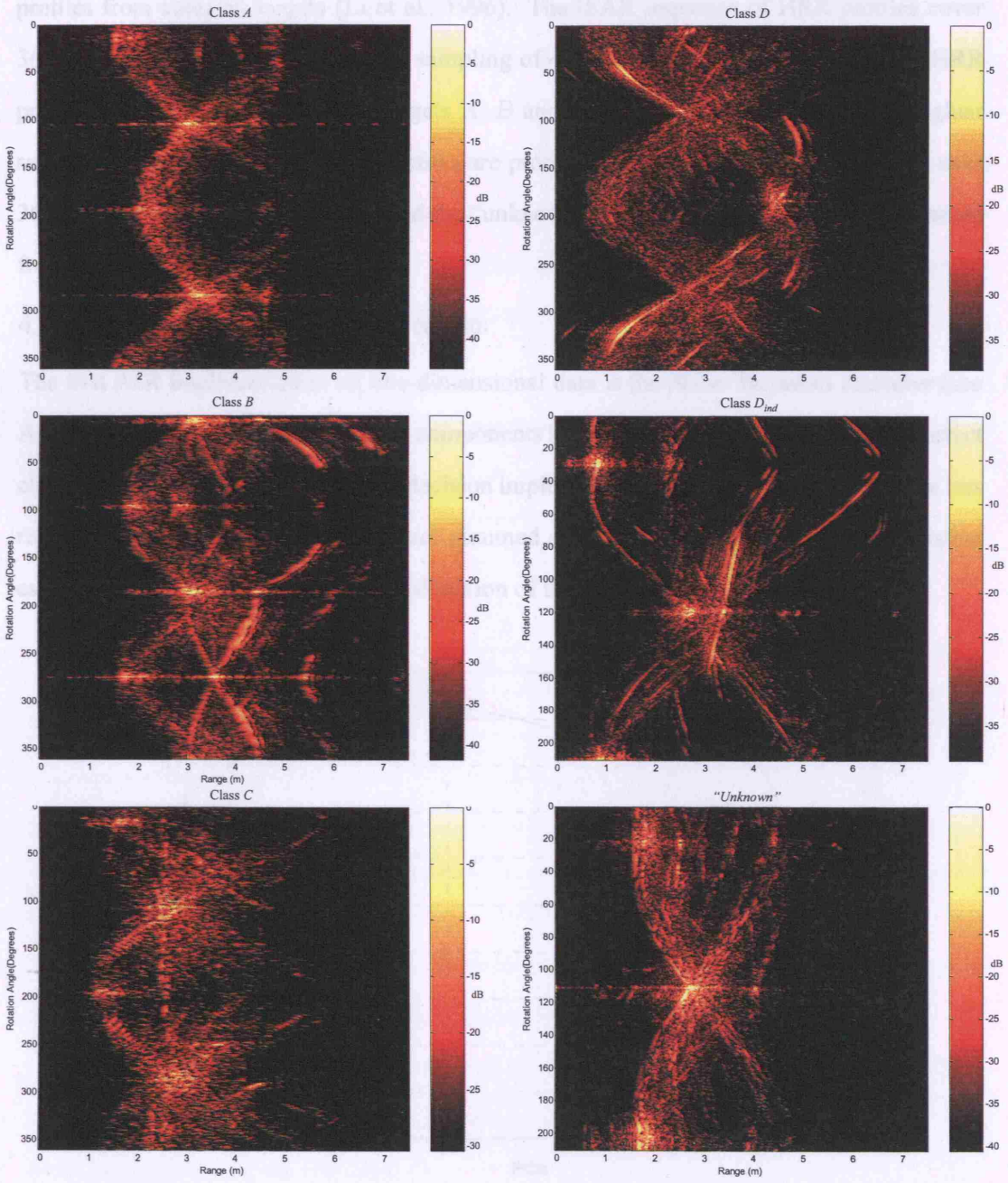

Figure 4.10: *History of HRR profiles over* 360° *(classes A, B, C, D) and* $\sim$ 210° *(D_{ind} and "Unknown").*

The former consists of 72 patterns, a range profile every 5 degrees of turntable rotation, whilst the latter contains the remaining profiles as a consequence of the analysis of the decorrelation over aspect angle, as expected from the correlation analysis for high resolution range profiles from complex targets (Li et al., 1996). The ISAR sequence of HRR profiles cover 360 degrees of rotation at the angular sampling of ~ 0.034 degrees. Therefore, 10495 HRR profiles are tested from the ADAS targets A, B and C; as regards the MIDAS data (angluar sampling ~ 0.028 degrees) 12672 profiles are processed from class D which is sensed over 360 degrees, while 7494 from D_{ind} and the "unknown" target class since they are illuminated from 217 degrees only.

4.4.3.1 Naïve Bayesian classifier results

The first ATR implementation on one-dimensional data is the Naïve Bayesian classifier (see Appendix A). The number of principal components has been chosen on the basis of the correct classification rate achieved in a forced decision implementation. Although it represents the less reliable case, the forced decision is here assumed as an environment providing a consistent estimate of the information content exploitation of the classifier.

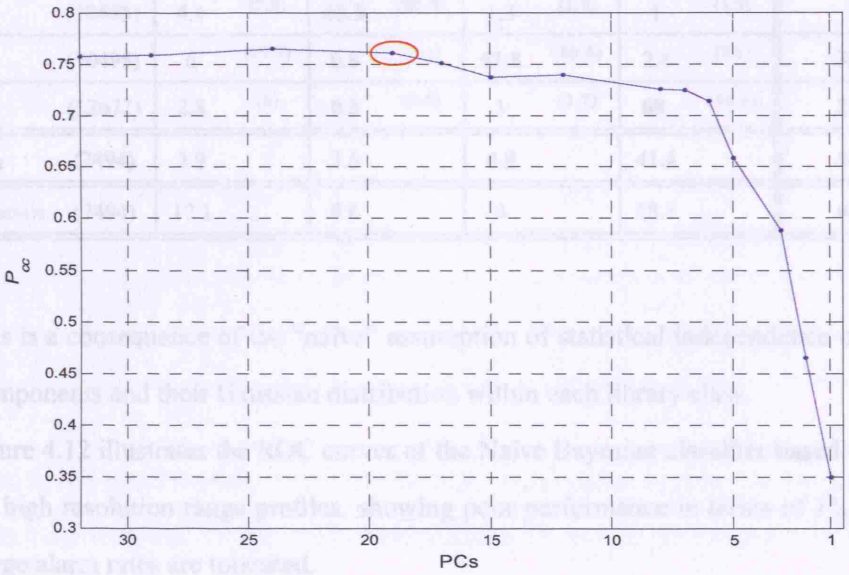

Figure 4.11: *Principal Components Analysis on high resolution range profiles ($\sim$ 8 cm) from complex target. The P_{cc} drops when less than 19 PCs are used as features.*

In figure 4.11, the probability of correct classification P_{cc} is plotted versus the number of principal components (PCs).

The original dimension of each range profile is 247 range cells, finally reduced to the most relevant 19 PCs that guarantee a classification accuracy comparable to the full elements of the range profile ($P_{cc} = 76.0\%$). The classifiers presented later in this section are also based on the same number of features.

Table 4.8 shows the performance of the Naïve Bayesian classifier after fixing the alarm rate at $P_{fa} = 39.6\%$ by setting the threshold on the posterior probability computed. This decision theoretic approach reveals a low declaration rate, although the generalisation capability is comparable to the probability of correct classification.

Table 4.8: *One-dimensional Naïve Bayesian classifier results results:* $P_{fa} = 39.6\%$, $P_d = 64.17\%$, $P_{cc} = 56.45\%$, $P_{gen} = 41.4\%$ *and* $R_{avg} = 87.57\%$.

		Output $^{(Reliability)}$ (%)								
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)	*Unknown*
A	(10495)	**42.7**	(76.4)	1.2	(1.8)	3.8	(6.6)	5.8	(7.5)	46.6
B	(10495)	4.4	(7.9)	**63.3**	(96.3)	1.3	(2.2)	1	(1.3)	30
C	(10495)	6	(10.7)	0.9	(1.4)	**51.8**	(89.5)	2.4	(3.1)	38.9
D	(12672)	2.8	(5)	0.3	(0.5)	1	(1.7)	**68**	(88.1)	27.9
D_{ind}	(7494)	3.9		3.5		4.9		**41.4**		46.3
Unknown	(7494)	17.1		0.4		4		18.1		60.4

This is a consequence of the "naïve" assumption of statistical independence of the principal components and their Gaussian distribution within each library class.

Figure 4.12 illustrates the ROC curves of the Naïve Bayesian classifier based on ADAS-MIDAS high resolution range profiles, showing poor performance in terms of P_{cc} and R_{avg} when large alarm rates are tolerated.

4.4.3.2 K-Nearest Neighbour Classifier Results

The classification performance is then measured using the K-Nearest Neighbour (K-NN, see Appendix B) classifier. The number of neighbors to be considered is crucial since a small K leads to noisy decision boundaries while a large K takes into account templates far away from the input one. Different values of K have been implemented, showing a maximum of correct classification rates in a forced decision scenario when $K = 5$ on a cross validation set of data (see Appendix B).

The rejection threshold is set on the minimum distance between the K nearest neighbours: if it exceeds a particular value d_{th} then the object is not declared. Depending on d_{th}, different false alarm rates can be achieved. A second threshold is also defined on the basis of the minimum number of votes necessary to declare a target. Since it is possible that the input has the same score for different classes (a draw), this needs to be thought of as another case of low degree of confidence and the target must be rejected.

Table 4.9: *One-dimensional K-NN classifier results:* $P_{fa} = 39.6\%$, $P_d = 69.7\%$, $P_{cc} = 63.05\%$, $P_{gen} = 63.5\%$ *and* $R_{avg} = 90.4\%$.

		Output $^{(Reliability)}$ (%)								
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)	*Unknown*
A	(10495)	**60.6**	(87.4)	3.4	(4.8)	1.6	(2.8)	2.5	(3)	31.9
B	(10495)	4.1	(5.9)	**64.9**	(91.7)	1.2	(2.2)	1.7	(2.1)	28.1
C	(10495)	3.1	(4.5)	1.4	(2)	**51.6**	(90.8)	2.6	(3.2)	41.3
D	(12672)	1.5	(2.2)	1.1	(1.5)	2.4	(4.2)	**75.1**	(91.7)	19.9
D_{ind}	(7494)	5.1		1.8		1.2		**63.5**		28.4
Unknown	(7494)	20.6		2.3		3.4		13.3		60.4

In table 4.9 the classification results for a 9-NN classifier after fixing the alarm rate at $P_{fa} = 39.6\%$ are shown. The generalisation performance are higher if compared to the Naïve Bayesian classifier. Besides, in figure 4.12, the ROC curves of the K-NN illustrate a higher reliability and correct classification rate at any false alarm probability if compared to the to the Naïve Bayesian classifier, although a deteriorated probability of correct classification for class

C, largely not declared.

4.4.3.3 Feed-forward Neural Networks Classifier Results

Finally, a Feed-forward Neural Networks (FANNs, see Appendix C) classifier is tested on the one-dimensional PCA processed patterns from ADAS and MIDAS targets.

The number of epochs and the learning rate of the network have been obtained on a series of measurements on the cross-validation set, useful to avoid overfitting problems (i.e. the network learns to recognise only the patterns used in the training phase, with the result of loosing the generalisation capabilities).

Table 4.10: *One-dimensional FANNs results:* $P_{fa} = 39.6\%$, $P_d = 72.92\%$, $P_{cc} = 69.87\%$, $P_{gen} = 49.8\%$ *and* $R_{avg} = 95.8\%$.

		Output $^{(Reliability)}$ (%)								
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)	*Unknown*
A	(10495)	**55.1**	(95.5)	0.6	(0.7)	2.3	(3.1)	2	(2.5)	40
B	(10495)	0.5	(0.9)	**79.7**	(98.4)	0.2	(0.3)	0.5	(0.6)	19.1
C	(10495)	1.7	(2.9)	0.6	(0.7)	**70.3**	(94.6)	1.7	(2.2)	25.7
D	(12672)	0.4	(0.7)	0.2	(0.2)	1.5	(2)	**74.4**	(94.7)	23.5
D_{ind}	(7494)	1.1		2.2		3.3		**49.8**		43.6
Unknown	(7494)	17		0.9		15.1		6.6		60.4

Table 4.10 shows the FANNs confusion matrix after fixing the alarm rate at $P_{fa} = 39.6\%$. The rejection threshold is set on the output layer of the network where each element represents the score for a particular class. The generalisation properties are lower than the K-NN classifier but higher reliability and correct classification rates are achieved. This can be deduced from the ROC curves of the three classifiers in figure 4.12.

The three classifiers exhibit different performance and distinct trends of ROC curves. The Neural networks present the larger Area Under the Curve (AUC) both in the $P_{fa} - P_d$ and $P_d - P_{cc}$ graphs. Furthermore, the FANNs classifier guarantees a low computational burden and will be subsequently employed when a large number of measurements is required.

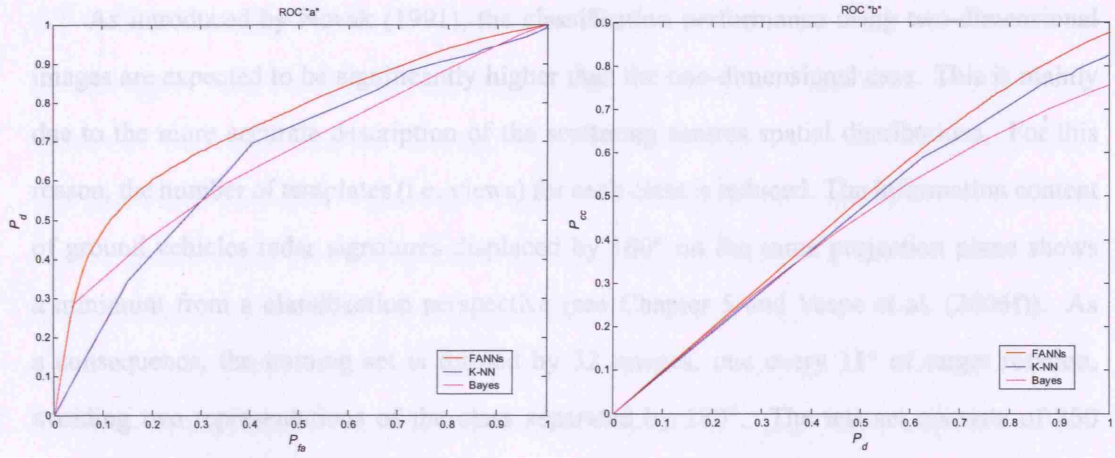

Figure 4.12: *One-Dimensional ROC curves from the Naïve Bayesian, K-NN and FANNs classifier.*

4.4.4 Two-Dimensional Results

In this section, the classification performance of a FANNs classifier only is described for the sake of brevity. The patterns to be classified are two-dimensional images from the ADAS-MIDAS targets. The range resolution synthesised is the same of the one-dimensional case ($\Delta r \sim 8$ cm), while after 2-D DFT processing an aperture $\psi = 2.83$ degrees, the cross range resolution achieved is: $\Delta r_c = 30$ cm.

Table 4.11: *Confusion matrix (single-aspect,* 8×30 *cm images):* $P_{fa} = 39.6\%$, $P_d = 95.7\%$, $P_{cc} = 93.9\%$. *Average reliability* $R_{avg} = 98.2\%$.

		Output $^{(Reliability)}$ (%)								
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)	*Unknown*
A	(350)	**85**	(98.7)	4.4	(4.3)	0.6	(0.6)	–	(–)	10
B	(350)	0.6	(0.7)	**98.3**	(95.7)	–	(–)	–	(–)	1.1
C	(350)	0.3	(0.3)	–	(–)	**97**	(99.1)	0.8	(0.9)	1.9
D	(350)	0.3	(0.3)	–	(–)	0.2	(0.3)	**95.3**	(99.1)	4.2
D_{ind}	(217)	2.7		12.9		–		**62.7**		21.7
Unknown	(217)	22.6		10.6		3.2		3.2		60.4

As introduced by Novak (1991), the classification performance using two-dimensional images are expected to be significantly higher than the one-dimensional case. This is mainly due to the more accurate description of the scattering centres spatial distributions. For this reason, the number of templates (i.e. views) for each class is reduced. The information content of ground vehicles radar signatures displaced by 180° on the same projection plane shows a minimum from a classification perspective (see Chapter 5 and Vespe et al. (2006f)). As a consequence, the training set is formed by 32 images, one every 11° of target rotation, avoiding two representations of the class separated by 180°. The test set consists of 350 images separated by $\sim$ 1°, neglecting the ones used to represent the target in the template library.

Table 4.11 shows the classification performance using two-dimensional images when $P_{fa} = 39.6\%$ is tolerated. The classifier performance is higher in terms of correct classification, generalisation, reliability and declaration rates if compared to the one-dimensional classifier. Nevertheless, forming a synthetic aperture image often requires a high computational burden and is not always possible since the dependence on the target movements. In figure 4.13, the ROC corresponding curves show a higher AUC if compared to the one-dimensional case, and therefore a more significant degree of trustworthiness since the reliability information can be derived from the slope of the ROC curve "b".

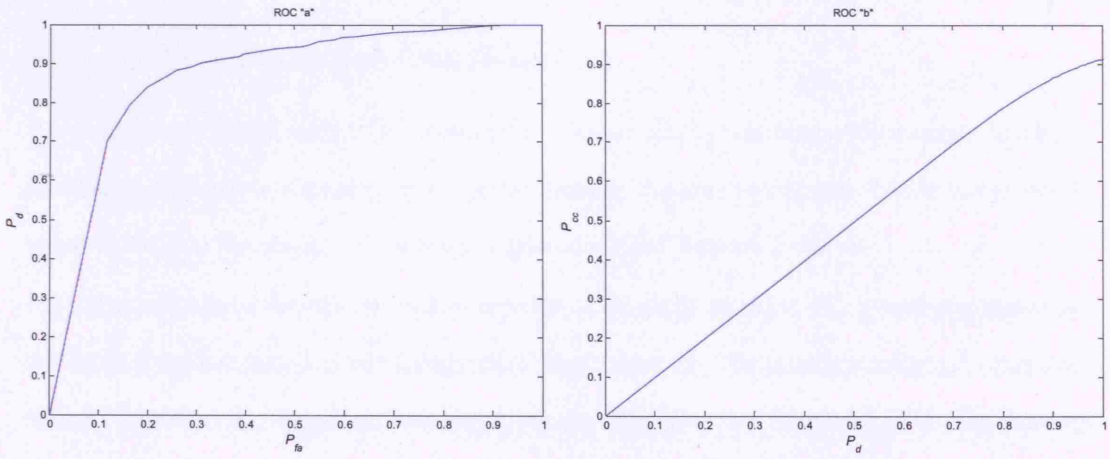

Figure 4.13: *Two-Dimensional ROC curves using a FANNs classifier.*

4.5 Summary

In this chapter, both the one- and two-dimensional ATR problems have been introduced and validated. They show also the weakness of such classification approaches especially when used in a real data scenario, where confusing and test variant targets may occur. The reliability and the generalisation capabilities clearly need to be further improved, and the classification results should not be relied on the forced decision implementation only. This, in many ways, provides the motivations for investigating alternatives to improve classification performance. A detailed examination of the data shows HRRP information to be clearly angle dependent. Thus it makes qualitative sense to exploit this using multiple perspectives.

The results reported here deal with a population set consisting of radar measurements from turntable experiments. A number of aspects are therefore responsible for the significant classification performance achieved, such as the very high SNR level, the low clutter return and the relative multipath, as well as the limited amount of classes. As a consequence, whilst encouraging, these conclusions should be treated with some caution as they are somewhat limited by the restricted available data and the particularly favourable scenario.

In the following chapters, perspective diversity is explored as a means of enhancing the target backscattering description and the ATR classification performance.

Chapter 5

Multi-Perspective Target Classification

In this chapter, after a brief overview of possible applications of Multi-Perspective ATR, the implementation of the classifiers introduced in Chapter 4 is described using multiple views of the target. Subsequently, the influence of the SNR on the classification performance is shown and its reduction using multiple looks at the target discussed. After explaining the non linear classification performance improvement with the number of perspectives used, an investigation on perspective location and displacement is presented using both one- and two-dimensional target signatures. Finally, the impact of cross range resolution is examined on ATR providing an insight into the image reconstruction techniques when the systems synthesises large apertures and, therefore, adjacent perspectives. In this first investigation, the classifier receives the different perspectives only, without the knowledge of their displacement.

5.1 Multi-Perspective Classification

We concentrate on the particular situation in which a single non-cooperative target has been previously detected and tracked by the radar system. Pre-processing raw data is necessary in order to increase the quality of the radar signatures (see Chapters 2 and 4).

The template library for each class consists of 72 range profiles. Each representative vector is taken approximately every 5 degrees of target rotation. The testing set of each class consists of the remaining range profiles excluding the templates (see Section 4.4.3). The features extracted after Principal Component Analysis (PCA) are the input attributes to the classifier.

Figure 5.1 shows a possible approximation of the multiple perspectives scenario: each

node of the network is assumed as having a fixed position as the target rotates by an angle ψ.

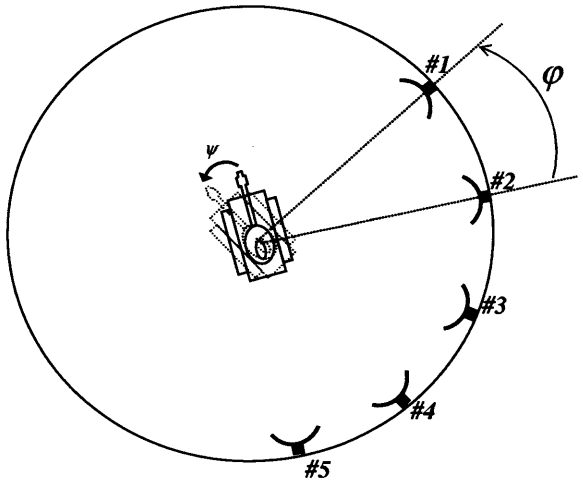

Figure 5.1: *Multi-Perspective environment deduced from an ISAR geometry.*

From the ADAS-MIDAS data (see Section 4.4.3), the target is on a turntable and all possible aspects (i.e. $\psi \in [0, 2\pi]$ are available, except for targets D_{ind} and the "unknown" class, sensed for a smaller range of orientations). The MSTAR data, described in Section 4.3.1, are collected from a SAR geometry following a circular trajectory around the target. The MSTAR image spatial sampling is shown in figure 5.2: the airborne radar flights over a circular aperture, collecting SAR images every $\sim 1°$ (blue dots).

The elevation angle δ can be assumed constant both for the template library (17 degrees) and the test set (15 degrees). The range is also maintained approximately constant so that the SNR does not deteriorate the quality of the backscattering description.

The perspective angle φ is the angular displacement between the lines-of-sight of two consecutive radar looks. From each of the radar positions either a series of range profiles can be generated as inputs to a one-dimensional classifier (ADAS-MIDAS data) or they can be processed into a synthetic aperture image (ADAS-MIDAS and MSTAR data) that can be the input to a two-dimensional classifier. It is therefore possible to perform classification using multiple perspectives. Equivalently to the traditional single-perspective performance

evaluation, where a classifier is tested for all the available headings in the test test, the M-P classifier is tested with all possible set of perspectives displaced by the particular topology under consideration.

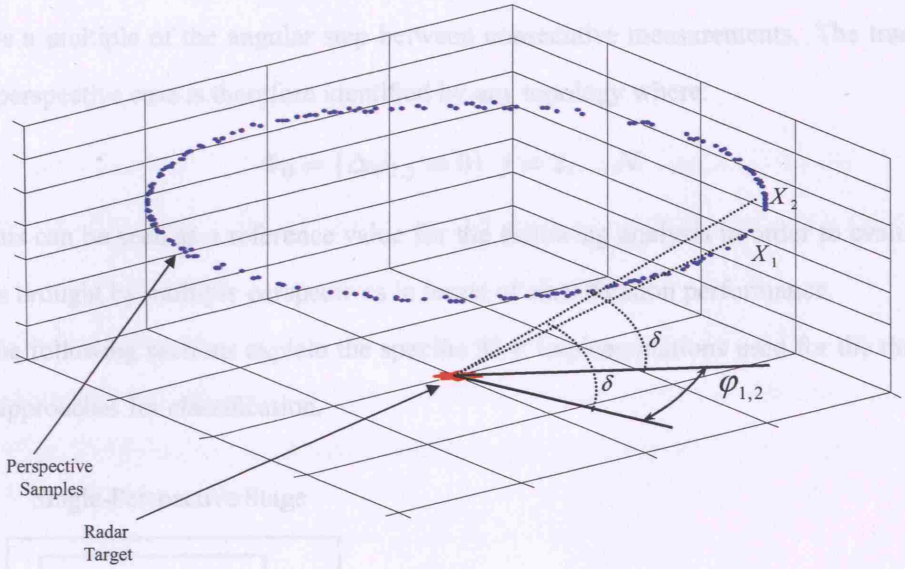

Figure 5.2: *MSTAR SAR space sampling from test set data.*

More formally, in an N-Perspective scenario, the parameter that distinguishes the perspective topology is the vector Φ:

$$\Phi = \{(\Delta\varphi; \Delta\delta)_{1,j}\} \;\; j = 2, ..., N \tag{5.1}$$

where the first node is taken as a reference, $\Delta\varphi_{1,j} = \varphi_j - \varphi_1$ is the angular displacement between the first and j-th nodes in azimuth, $\Delta\delta_{1,j} = \delta_j - \delta_1$ is the displacement in elevation and N is the number of perspectives. In this work we limit the discussion to a two-dimensional problem, considering multiple perspectives in the azimuth domain only:

$$\Phi = \{\Delta\varphi_{1,j}\} \;\; j = 2, ..., N \tag{5.2}$$

The index j is then considered ranging from 2 to N for symmetry, since it is effective their absolute displacement on the information content of the multiple signatures. For example in a two-Perspective (2-P) scenario, the parameter that distinguishes the perspective node locations

is their relative angular displacement $\Delta\varphi_{1,2} = \varphi_2 - \varphi_1$. Hence, after fixing $\Delta\varphi_{1,2}$, the 2-P classifier is tested with all possible pairs of signatures $\{X_1, X_2\}$ displaced by that angle covering all the possible orientations of the target. Having a test set consisting of M profiles, the same number of pairs can be formed to test the 2-P classifier for a given topology, provided that it is a multiple of the angular step between consecutive measurements. The traditional single-perspective case is therefore identified by any topology where:

$$\Phi_0 = \{\Delta\varphi_{1,j} = 0\} \quad j = 2, ..., N \tag{5.3}$$

This can be seen as a reference value for the following analysis in order to evaluate the benefits brought by multiple perspectives in terms of classification performance.

The following sections explain the specific M-P implementations used for the three different approaches for classification.

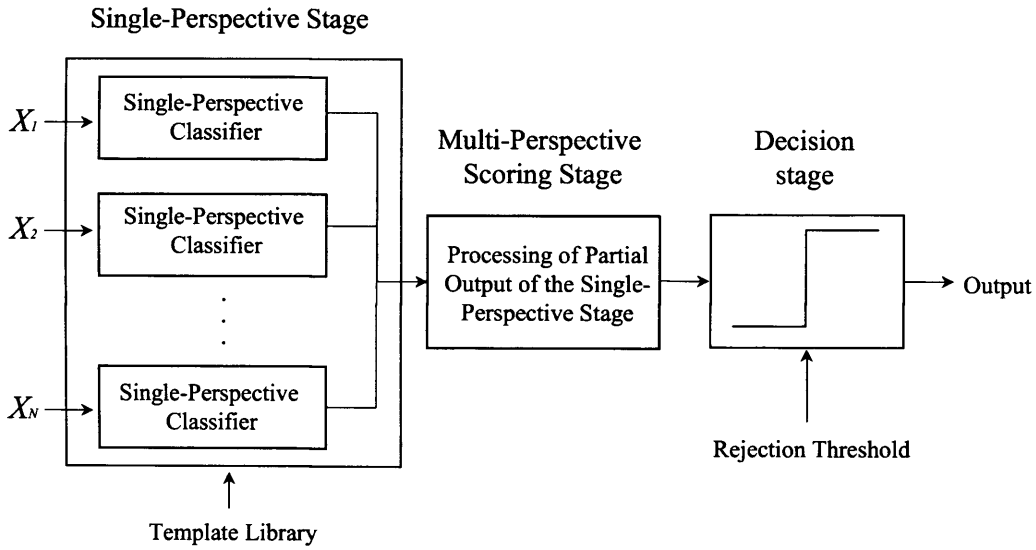

Figure 5.3: *The Multi-Perspective classifier. A first Single-Perspective stage enables the parallelization of the process: the nodes independently produce the partial outputs that are subsequently processed by the M-P stage in order to eventually produce a score.*

A possible M-P classification methodology is illustrated in figure 5.3. The classification block producing a score to be compared with the rejection threshold (see figure 4.1) is subdivided into two stages: the first is a single-perspective stage, where the perspectives $X_1, ..., X_N$

are partially processed to generate a first score. The second is the M-P stage, where the single-perspective outputs are processed. Finally, after combining the partial outcomes in the M-P stage a score is eventually produced.

This structure allows for parallel implementation of the procedure and reducing the time of execution.

5.1.1 Multi-Perspective Bayesian Classifier

The concepts discussed in Chapter 4 and Appendix A are now integrated for the M-P Naïve Bayesian classifier. Here we consider a network of N radars, or more generally a set of N perspectives. The sequence of one- or two-dimensional signatures $\{X_j : j = 1, ..., N\}$ represent the information collected by the system. The M-P posterior probability $P(C_i \mid X_1, ..., X_N)$ of the sequence of perspectives conditioned to C_i is the probability that the sequence belongs to that class. Firstly the single-perspective posterior probabilities $P(C_i \mid X_j)$ are computed:

$$P(C_i \mid X_j) = \frac{P(X_j \mid C_i)P(C_i)}{\sum_{k=1}^{n_c} P(X_i \mid C_k)P(C_k)} \tag{5.4}$$

Subsequently, assuming that the perspectives are statistically independent, they are combined producing the M-P posterior probability using the principle of Bayesian updating (Baker and Boult, 1990):

$$P(C_i \mid X_1, ..., X_N) = \frac{\prod_{j=1}^{N} P(C_i \mid X_j)}{\sum_{i=1}^{n_c} \left[\prod_{j=1}^{N} P(C_i \mid X_j)\right]} \tag{5.5}$$

where the denominator is a normalisation factor. Equation 5.5 gives an estimate of the posterior probability that the sequence of perspectives has been generated by a particular class C_i. As explained in Chapter 4, this estimate is used as a score for C_i and a rejection threshold can be applied.

5.1.2 Multi-Perspective K-NN Classifier

There are three stages of the M-P K-NN classifier (see Appendix B) which are implemented as follows (figure 5.4):

1. The single-perspective stage: after the collection of the sequence of feature vectors $\{X_j : j = 1, ..., N\}$ where N is the number of sensed signatures, the same number of

single-perspective classifiers is implemented. The j-th classifier computes a vector D_j consisting of the K minimum distances from the templates.

2. M-P processing: the set of vectors $\{D_j : j = 1, ..., N\}$ is processed and the minimum K distances are again selected giving a weight for the decision.

3. The rejection threshold, based on a minimum distance tolerated to make a decision (outlier detection), is applied to the Multi-Perspective score. A further threshold needs to be taken into account when more than a single class have the same number of votes. In this case, the target should be declared but not assigned to any class.

4. Classification: the input sequence of feature vectors is associated with the class with the greatest number of weights.

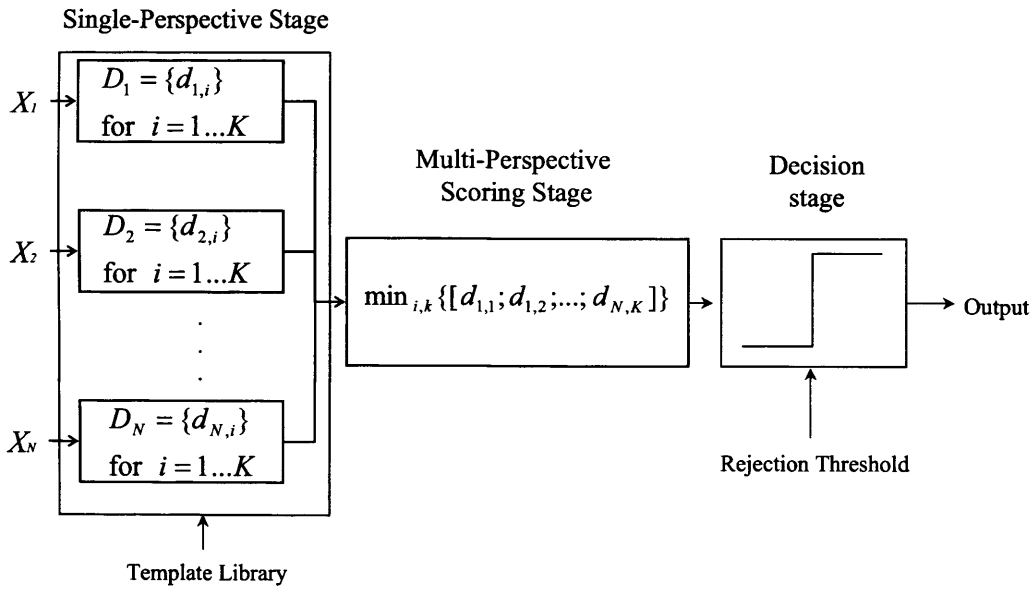

Figure 5.4: *Multi-Perspective K-NN classifier.*

5.1.3 Multi-Perspective Classifier using Neural Networks

As described in Appendix C, the output Y_i of a Neural Network classifier having X as input can be thought of as an estimate of the posterior probability $P(C_i \mid X)$. As a consequence, the single-perspective stage for the signature X_j represents a partial score distribution for

the set of n_c classes. Their combination using equation 5.5 would therefore provide for the Multi-Perspective stage output (figure 5.3). In figure 5.5, this structure is clearly visible: every network is trained on the template set described by equation 4.1.

As for the Naïve Bayesian classifier, the rejection threshold is a requirement on the minimum posterior probability.

This approach to Multi-Perspective processing has been adopted for both Feed-Forward Neural Networks (FANNs) and Radial Basis Function Neural Networks (RBFNNs) classifier, using therefore both ADAS-MIDAS and MSTAR data, and their performance is discussed in the following sections in terms of probability of correct classification (number of perspective and SNR, optimal perspective location), ROC curves, reliability and generalisation capabilities.

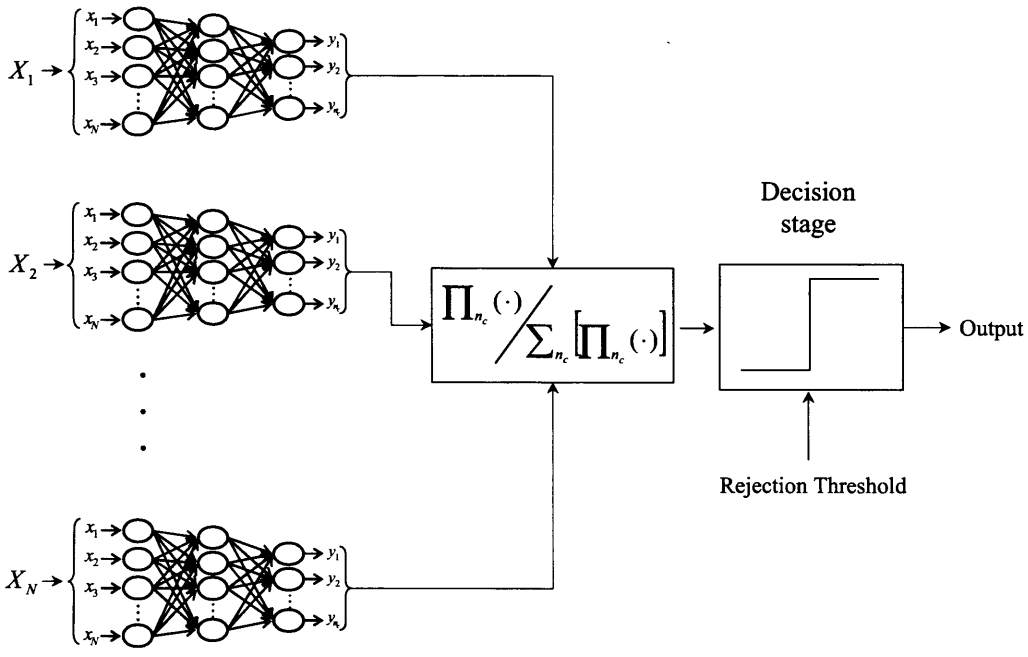

Figure 5.5: *Multi-Perspective FANNs classifier implementation.*

5.1.4 ADAS-MIDAS data: Multiple Perspectives and Probability of Correct Classification

The analysis of the benefits brought by multiple perspective in terms of classification performance is first focused on the increase of the correct classification rates in a forced decision environment.

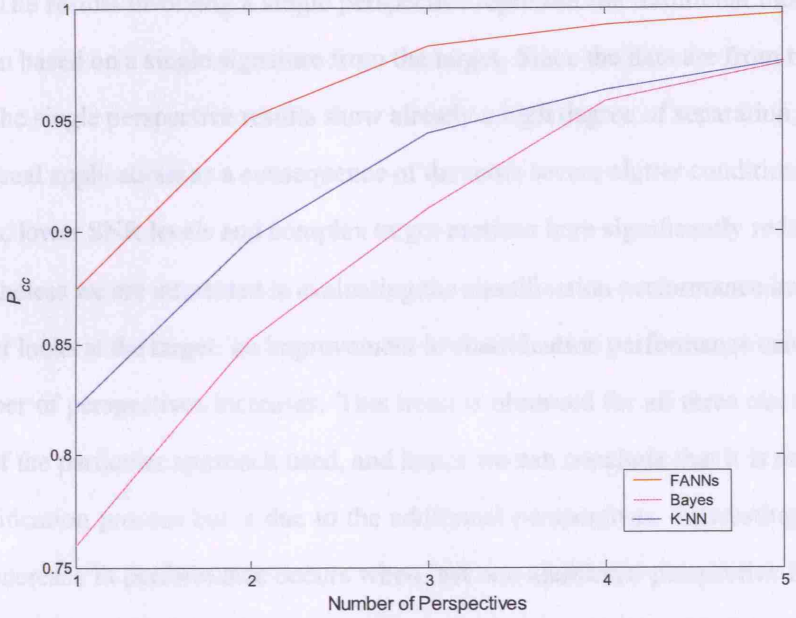

Figure 5.6: *Probability of correct classification of three pattern recognition approaches (K-NN, Bayesian and FANNs classifiers) implemented using multiple perspectives where the angular displacement between successive looks is* $\Delta\varphi_{i,i+1} = 20°$.

A first perspective topology is assumed to present a progressive number of looks equally displaced by 20°: firstly, the two-perspective case is considered and identified by the topology $\Phi_{2P} = \{\Delta\varphi_{1,2} = 20°\}$. Then the three-perspective case is examined by adding $\varphi_{1,3} = 40°$ to the topology, i.e. $(\Phi_{3P} = \{20°, 40°\})$, the four-perspective $(\Phi_{4P} = \{20°, 40°, 60°\})$, and finally the five perspective scenario is investigated $(\Phi_{5P} = \{20°, 40°, 60°, 80°\})$. This angular displacement between perspectives has been chosen from a preliminary analysis of the range profiles in the dataset, which exhibit a significant degree of decorrelation after a few degrees

of aspect change. Therefore, 20° guarantees a high degree of diversification in the information contained in the multiple perspectives. This is discussed in detail in the following section.

The ADAS-MIDAS data are used in form of $\sim$ 8 cm resolution range profiles and the Multi-Perspective P_{cc} performance evaluated for the classifiers previously described, using the library template set discussed in Chapter 4.

Figure 5.6 shows the probability of correct classification versus the number of perspectives used. The results involving a single perspective represent the traditional monostatic implementation based on a single signature from the target. Since the data are from turntable experiments, the single perspective results show already a high degree of separation, not usually achieved in real applications as a consequence of the more severe clutter conditions, multipath components, lower SNR levels and complex target motions here significantly reduced.

Nevertheless we are interested in evaluating the classification performance improvements using further looks at the target: an improvement in classification performance can be deduced as the number of perspectives increases. This trend is observed for all three classifiers, independently of the particular approach used, and hence we can conclude that it is not a function of the classification process but is due to the additional perspectives. Interestingly, the most significant increase in performance occurs when just one additional perspective is employed. Furthermore, it emphasises that the M-P improvements are not linear with the number of perspectives involved. On average a 8.1% improvement in the correct classification rate, with respect to the single perspective classifier, is reached when two perspectives are used while the benefit in relation to the 2-P classifier is 4.6% when three perspectives are used. It is 2.5% when going from three to four perspectives and lastly 1.1% in going from four to five perspectives. This is a consequence of those signatures that relate to particular target orientations where specular reflections occur and the profile is thus dominated by only a very few scattering centres. These patterns are common to all the target classes and consequently are more easily misclassified. Nevertheless, for these target orientations, the decision confidence of the classifier is low and is therefore significantly aided by a second perspective. On the other hand, if both the first and second perspectives correspond to those signatures presenting reduced information, a third perspective is less effective in improving performance.

A third perspective has less influence to correct the weights of the others and thus the improvement from subsequent perspectives is progressively attenuated. This trend may be different if the perspective displacement is processed by the classifier as useful information, as discussed in Chapter 6.

The non linear relationship between P_{cc} and the number of perspectives suggests that only the employment of a small number of looks is sufficient to achieve a significant improvement, relatively reduced when using further perspectives. This may be important in designing a network of radars, since the number of nodes increases with the system complexity and cost. In this case a limited number of nodes is necessary in order to achieve the best trade-off between system complexity and classification performance improvement.

The FANNs classifier shows the best correct classification performance with the increase of the number of perspectives as well as the minimum computational burden. As a consequence, multilayer neural networks have been selected as representative classifiers for the following sections.

5.2 ADAS Data: Multiple Perspectives vs SNR

The extent to which Signal to Noise Ratio (SNR) affects classification and whether Multi-Perspective scenarios are effective at different SNR levels are now examined. The FANN classifier has been used for this particular task. The range profiles I and Q components are corrupted with additive white Gaussian noise. The SNR of the data is measured as the distance in dB from the maximum magnitude in the HRR profile belonging to a particular target, and the noise level after removing the ZDC component (see Section 2.1.2.1).

From each target class C_i, the noise level is measured in the non-target area before Target Normalisation (see Chapter 2). The target signal level is estimated as the mean value of the maximum target backscattering over the range profile history covering 360 degrees headings, and the SNR_i estimated as follows:

$$\text{SNR}_i = \frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} 20 \log \left(\frac{r_{m,k}}{\mu_m} \right) \tag{5.6}$$

where μ_n is the mean noise intensity measured in the m-th k-dimensional range profile

$RP_m = (r_{m,1}, r_{m,2}, ..., r_{m,k})$, and M is the number of profiles collected from the i-th target class. The result of applying this procedure to the ADAS and MIDAS data leads to a measured SNR $= 1/n_c \sum_i \text{SNR}_i = 27.8$ dB on the original data. Subsequently, the classifier is tested with range profiles presenting a progressively lower SNR. Then, the SNR is measured as above, and classification attempted on PCs extracted from the one-dimensional signatures. The original training set (i.e. not corrupted) is preserved, as well as the Neural Network weights previously obtained for the classification experiment in Section 4.4.3.3.

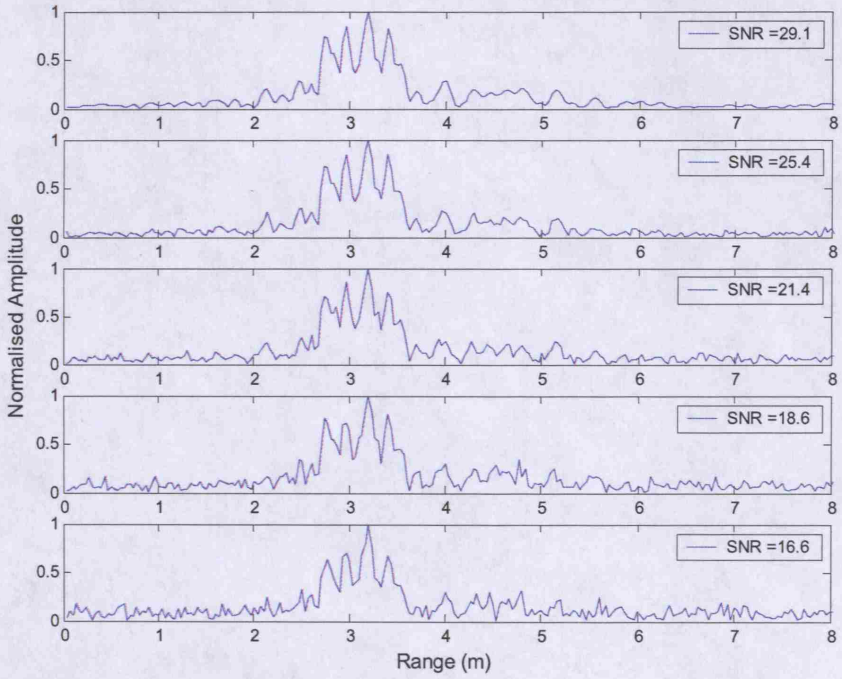

Figure 5.7: *Range profile belonging to class C progressively corrupted with noise.*

In figure 5.7, an example of a single range profile from class C is plotted at different SNR levels. The rotation angle is 110 degrees (see figure 4.10) and the corresponding radar length is approximately four metres, spanning the range bins from 2 m to 6 m. Some scatterers are occluded because of the low grazing angle, therefore the radar length does not correspond to the actual physical length of the target.

As can be observed, the low-level scatterers of the target gradually disappear in noise. In addition the peak values become more and more corrupted with the result of reducing the

useful information in terms of the higher peak values and their relative locations. Furthermore, the radar length of the target is no longer deducible with any great certainty. Clearly this will have an adverse effect on classification performance no matter how many perspectives are used.

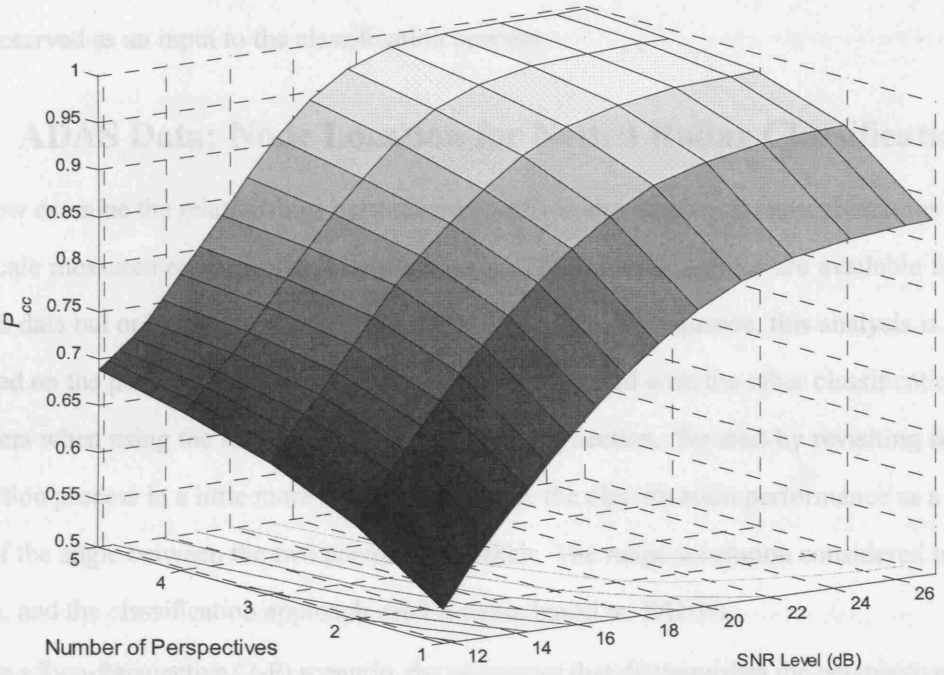

Figure 5.8: *Probability of correct classification for a Multi-Perspective FANNs classifier tested at different SNRs (1-D range profiles, $\Delta r \sim 8$ cm, four-class problem, $\Delta\varphi_{i,i+1} = 20°$).*

In figure 5.8 the performance of the FANN classifier only is plotted versus the number of perspectives used and SNR levels, showing how the enhancement in classification varies with different noise levels. The graph illustrates an increase in classification performance with numbers of perspectives in each case, particularly valuable for lowest SNR levels. However, below an SNR of 17 dB the performance quickly degrades indicating that classifiers will be upset by relatively small amounts of noise.

This is because of the information loss caused by the increasing noise that corrupts the range profile. The relatively high SNRs at which this occurs can be partly explained by the fact this it is a combination of smaller magnitude and larger magnitude scatterers that provide

the classification information. The smaller magnitude scatterers are more severely affected by the increasing noise and hence quickly have an impact on reducing performance. As the number of perspectives increases, the information is partially recovered with the help of other perspectives since the noise is not correlated and hence modifies the profiles differently. This result highlights the importance of having adequate SNR, such that small magnitude scatterers are preserved as an input to the classification process.

5.3 ADAS Data: Node Location for Netted Radar Classification

We now examine the relationships between perspectives and target geometrical features using full scale measurements of representative targets. The full 360 degrees are available for the ADAS data but only for class D of the MIDAS data. As a consequence, this analysis is again focused on the probability of correct classification, integrated with the other classification parameters when using the MSTAR data in the following section. We start by revisiting feature extraction process in a little more detail and examine the classification performance as a function of the angle between the two perspectives taken. The range resolution considered here is 30 cm, and the classification approach used is again based on FANNs.

In a Two-Perspective (2-P) scenario, the parameter that distinguishes the perspective node locations is their relative angular displacement $\Delta\varphi_{1,2} = \varphi_2 - \varphi_1$. The training set of representative vectors for each class is made up by 36 range profiles, taken approximately every 10° degrees of target rotation. The testing set of each class consists of the remaining range profiles neglecting the templates. The M-P classifier can be seen as the combination of N single-perspective classifiers whereas the eventual decision is made by processing the outputs of each FANN. For this first stage of investigation, the angle $\Delta\varphi_{1,2}$ is not processed as information by the M-P classifier.

Features from radar signatures have been extracted in order to reduce the intrinsic redundancy of data and simplify the model. This methodology, to a certain extent, decreases the overall correct classification rates but, on the other hand, makes the representation model in the feature space less complex yielding a classification process consistent with those features effectively characterising the object. A typical HRR profile is shown in figure 5.9, where a

threshold is applied to the HRR profile. The threshold is determined by measuring the mean intensity value neglecting the maximum peak, after normalising the profile. This guarantees adaptive target area isolation and less sensitivy to the main scatterer reflection.

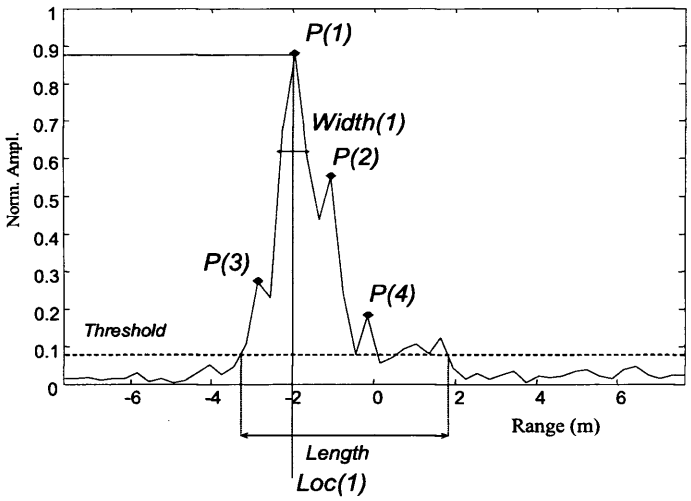

Figure 5.9: *Feature extraction from HRR profile based on peaks locations and intensities.*

Therefore, the radar length of the target for that particular orientation is measured as the distance between the first and last threshold crossings. This is the first component of the feature vector f. The second component is a measure of the average backscattering of the target whilst the successive M triples contain the information of the M peaks extracted in terms of amplitude, location and width. If the number P of peaks above the threshold is less than M, the last $M - P$ triples are set to zero (Huaitie et al., 1997). Different numbers of peaks have been extracted until the classification process revealed a certain degree of robustness. For $M = 4$, the feature vector has a dimension of 14 elements while 52 range bin values make up the raw echo profile. As a consequence of this feature extraction algorithm, the information of the low level scatterers approximately below $14dB$ might be not used because of the presence of a number of more intense returns from particular target headings. The number M is therefore crucial and has to be estimated depending on the resolution, the nature of the target and the SNR of the data.

Subsequently, PCA has been applied to the feature vectors and, by choosing the largest

eigenvectors from the original 14 of the covariance matrix, the dimension is additionally reduced. In table 5.1 the resulting confusion matrix for $PC = 10$ is shown for the four-class population problem using a single perspective of the target.

Radar classification is highly dependent on the object orientation. Since radar targets are usually manmade objects, they often present a number of 3-D symmetries. In this work, a low grazing angle is used to simulate a ground-based scenario, where the radar system, the target and their relative motion vectors lie on the same plane.

Table 5.1: *Single-perspective FANN confusion matrix on features after PCA using* $P = 10$ *(*$P_{cc} = 74.4\%$*, average reliability* $R_{avg} = 75.1\%$*).*

		Output $^{(Reliability)}$ (%)							
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)
A	(10495)	**75.8**	(78.8)	11.2	(10)	9.6	(8.8)	3.4	(4.1)
B	(10495)	4.7	(4.9)	**80.9**	(72.3)	13.3	(12.3)	1.1	(1.3)
C	(10495)	2.8	(2.9)	19	(17)	**69.9**	(64.7)	8.3	(9.9)
D	(12672)	12.9	(13.4)	0.7	(0.7)	15.3	(14.2)	**71.1**	(84.7)

This geometry allows us to consider the 2-D problem, whereby the perspectives represented by 1-D signatures collected by the measurement system are on the same plane. For example, for most of the 2-D ground-vehicle orientations, with 180° between the two perspectives, the corresponding profiles might expect to be quite highly correlated. This is due to the 180° symmetry typically exhibited by vehicles and hence little extra information is added. If this is the case it will cause a reduction in target characterisation and eventually, of M-P probability of correct classification improvement. However, details such as rear-view mirrors, the antenna position and any non-symmetrical or moving parts will change the two signatures producing benefits on P_{cc} when compared with single-perspective classifiers.

We now consider the relationship between nodes location and target orientation, investigated for the cases of both two and three-perspective classifiers. In a 2-P scenario the angular

perspective displacement between radars is the discriminant factor for the combined P_{cc}: as the two range profiles decorrelate (i.e. $\Delta\varphi_{1,2}$ increases) the information content of the pair-increases.

Figure 5.10 shows the classification performance as a function of angular separation of the two perspectives. The equivalent monostatic P_{cc} is shown at the 0° and 360° positions.

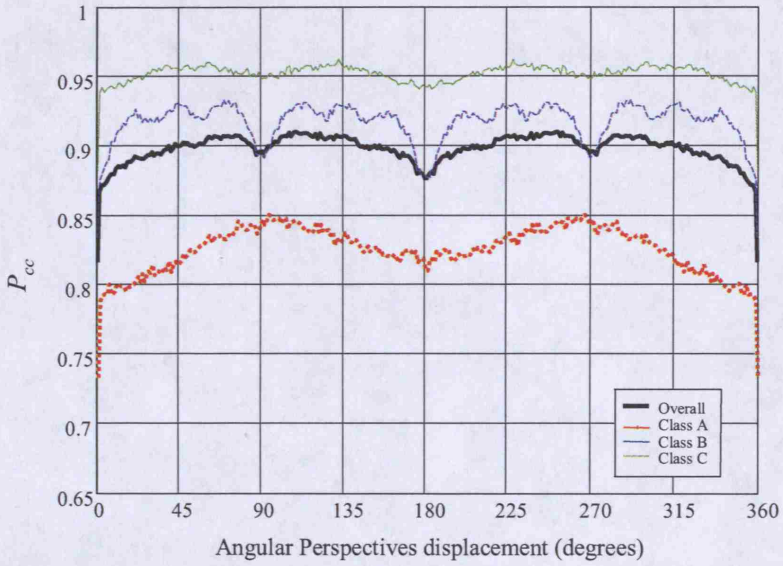

Figure 5.10: *2-P correct classification rates versus the angular displacement $\Delta\varphi_{1,2}$ for the three-class problem.*

The single target classification rates show how the global accuracy depends on the peculiar geometric features of the target. The drop at $\Delta\varphi_{1,2} = 180°$ (as previously hypothesised) is due to the multiple axes of symmetries of the targets and it is visible for all the classes. For targets A and B there is also a drop at 45° indicating a possible further degree of symmetry.

The relationship between M-P classification and range profiles information content can be deduced from figures 5.11, 5.12 and 5.13 where the cross-correlation between two profiles X_1 and X_2 collected from different perspectives is represented for the three classes A, B and C respectively. The maximum values are in correspondence to the auto-correlation lines $(X_1 = X_2)$, meaning that the information content of the pair is highly correlated and therefore

almost equivalent to the perspectives taken individually (mono-perspective case).

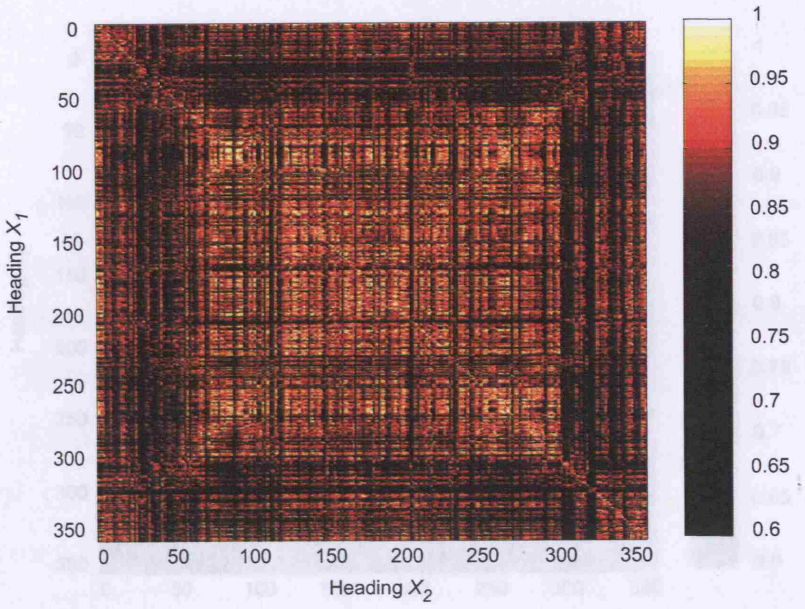

Figure 5.11: *Cross-correlation between profiles belonging to class A.*

As a consequence, high cross-correlation between patterns means that the M-P environment yields performance close to that of the single-perspective case. Conversely, a low cross-correlation indicates more disparate information although it does not always lead to a more accurate target description. The regions of high cross-correlation influence the M-P classifier: when 90° of separation occurs, the profiles between two perspectives of class B (figure 5.12) taken in the range [120-150] degrees are highly correlated with the ones belonging to the orientations [210-240] degrees. This can now be linked to the significant drop of probability of correct classification in figure 5.10, occurring for this class when the perspective separation between looks is $\pi/2$, π. Other regions showing high cross-correlation are therefore responsible for the drops of correct classification rates in the neighbourhood of $\pi/6$, $2\pi/3$ and so forth.

If we analyse the target of class C, a lower cross-correlation can be observed in regions where one of the signatures is in the neighbourhood of $\{\pi/2, \pi, 3\pi/2\}$ and the second one is $\pi/2$ or π away. This means that the information content of such perspective displacements is

expected to be close to the mono-perspective case, and this is confirmed by the class C trend shown in figure 5.10.

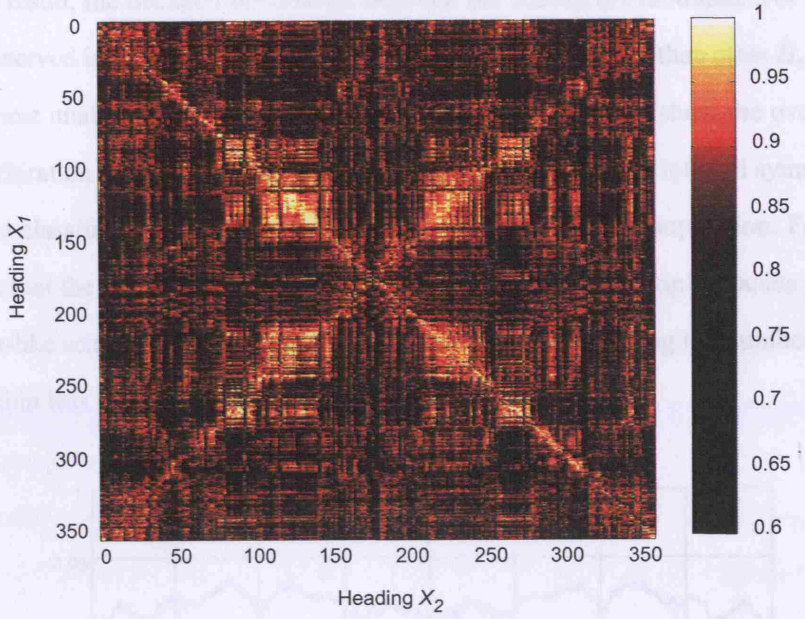

Figure 5.12: *Cross-correlation between profiles belonging to class B.*

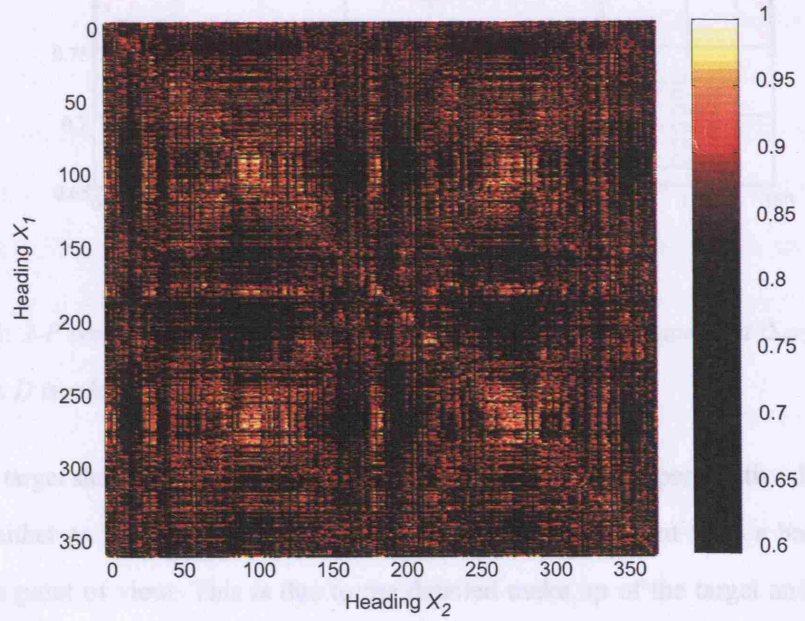

Figure 5.13: *Cross-correlation between profiles belonging to class C.*

We now add to the dataset another sub-population class D and a new M-P FANN classifier. All the internal parameters of the Neural Network were changed by the new learning phase. As a result, the decision boundaries between the classes are modified. For this reason, as can be observed in figure 5.14, class C is more likely misclassified than class B, whose P_{cc} remains almost unaltered. The four-class correct classification rates show the overall performance deterioration in terms of classification. On the other hand, the internal symmetries that influence the classifier remain unchanged by adding elements to the population. For example, it is thought that the target belonging to class B has a number of multiple-bounce phenomena from corner-like scatterers. Their persistence is less than 15°, causing the number of relative maxima within less than 90° of separation.

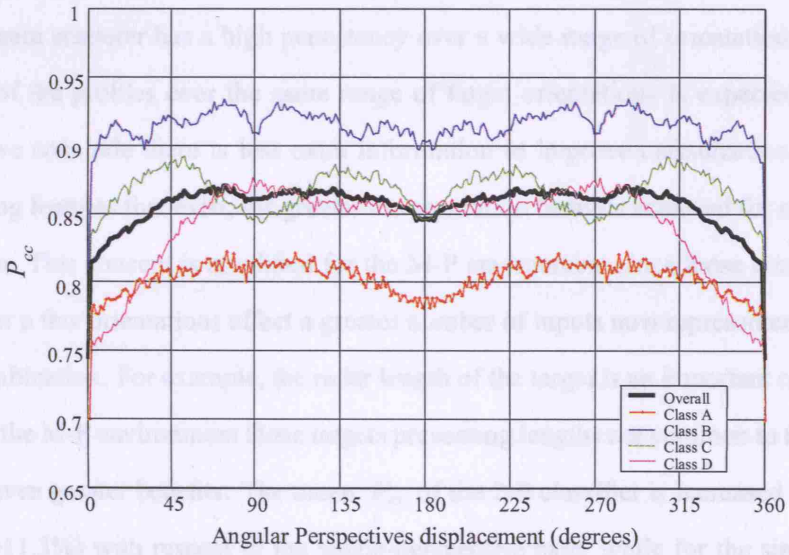

Figure 5.14: *2-P correct classification rates versus the angular displacement $\Delta\varphi_{1,2}$ when the fourth class D is added.*

Every target shows different trends except for 180°: 90° angular perspective displacement can mean either an improvement or reduction in information content from a backscattering information point of view. This is due to the detailed make up of the target and the way in which differential symmetries can be exhibited. This can make the overall M-P classification rates less sensitive to the angular perspective displacement of the nodes.

The effects of geometrical symmetries on CCR performance can also be seen in the neighbourhood of $\Delta\varphi_{1,2} = 90°$ and $\Delta\varphi_{1,2} = 270°$. This is verified when the relationship between two perspective displacements $\Delta\varphi_{1,2}$ and $\Delta\varphi'_{1,2}$ is:

$$\Delta\varphi_{1,2} = \pi - \Delta\varphi'_{1,2} \tag{5.7}$$

When the perspective condition expressed in the above equation is verified, the 2-P performance is similar because of the intrinsic geometrical symmetries of classes A, B, and C. This can be observed in both the three and four-class problems shown in figures 5.10 and 5.14, illustrating the M-P independence of the cross-affinity between different targets. The decorrelation rate of the HRR profiles also seems to vary depending on the particular target. In general, it appears proportional to the number of wave trapping features and their persistency. If a single main scatterer has a high persistency over a wide range of orientations, the cross-correlation of the profiles over the same range of target orientations is expected to be high and hence we conclude there is less extra information to improve classification. The more distinguishing features that exist, the greater the separation benefits achieved for single-aspect classification. This concept is amplified for the M-P environment since those crucial features appearing for a few orientations affect a greater number of inputs now represented by the perspective combination. For example, the radar length of the target is an important classification feature. For the M-P environment those targets presenting lengths not common to the other objects it has even greater benefits. The mean P_{cc} of the 2-P classifier is increased from 74.4% to 85.7% (+11.3%) with respect to the single-perspective case, while for the single class D the probability of correct classification increases by +13.1%. However, this result is highly averaged but is indicative of the overall performance improvement.

In figure 5.15 the classification rates for three perspectives are shown. Here the first perspective is fixed whilst the other two slide around the target covering all 360°. Thus these are shown as a function of the angular displacement $\Delta\varphi_{1,2}$ between the second and the first perspective, and $\Delta\varphi_{1,3}$ between the third and the first perspective.

The origin represents the mono-perspective case and again indicates the overall improvement offered by the Multi-Perspective approach. The bisector line corresponds to two radar

systems at the same location, while the lines parallel to the bisector symbolise two perspectives displaced by 90° and 180°. The inherent symmetry in the radar signature of the vehicle gives rise to the relatively regular structure shown in figure 5.15.

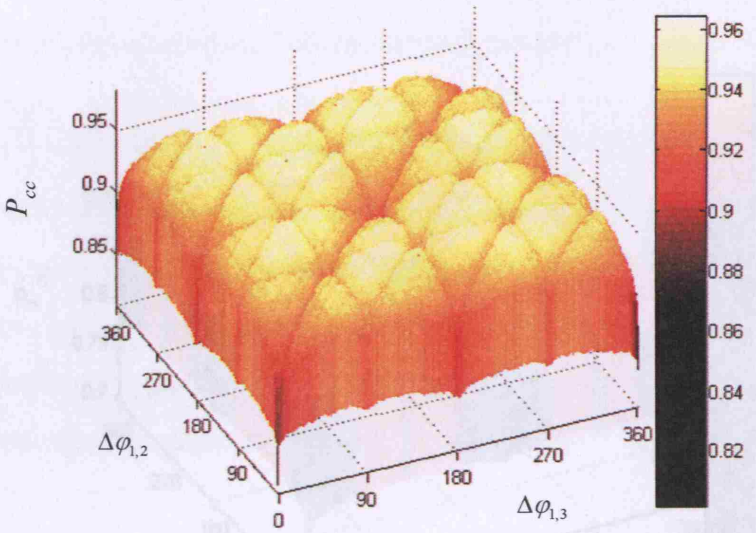

Figure 5.15: *Three-Perspective P_{cc} versus the angular displacements $\Delta\varphi_{1,2}$ and $\Delta\varphi_{1,3}$ for the three-class problem.*

In a 2-Perspective scenario, if the two nodes view the target from the same perspective (i.e. the two radars have the same LOS) then this gives the mono-perspective classification performance. This is not the case for a 3-Perspective network. This is a consequence of weighting twice the perspective of two nodes and once the third node's perspective. As a result, the 3-Perspective performance when $\Delta\varphi_{1,2} = 0$ and $\Delta\varphi_{1,3} \neq 0$ is worse than the 2-Perspective scenario that simply neglects one of the coinciding perspectives.

In figure 5.16, the CCR of the four-class population problem (the target class D has been introduced) is represented with respect to the angular displacements between the three perspectives. As it has been observed for the 2-Perspective case, the classification performance is less sensitive to the aspect angle when considering a greater number of classes problem. This may suggest that in a real environment with a large number of classes, the M-P classifica-

tion could be equally effective for any target position provided that each node LOS is spaced enough from the others in order to collect uncorrelated signatures. The 3-P classifier probability of correct classification is increased from 74.4% to 90.0% (+15.6%) when compared to the single-perspective case.

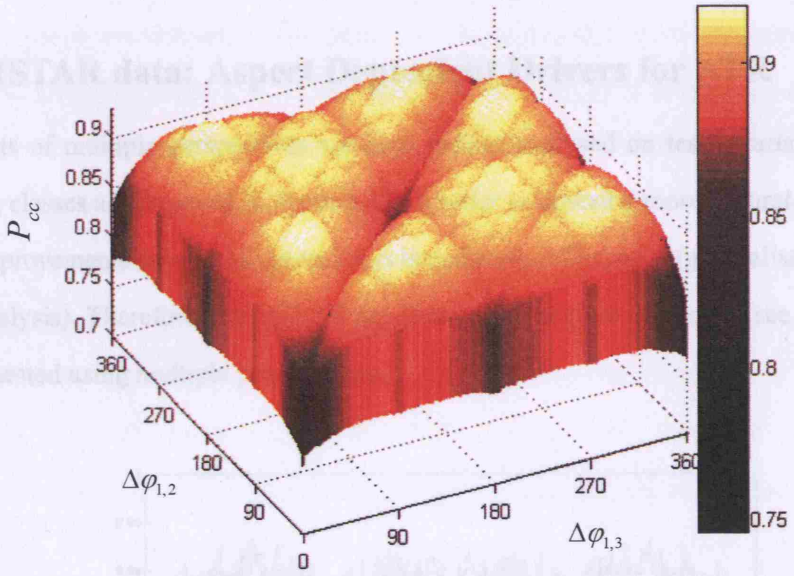

Figure 5.16: *Three-Perspective Probability of Correct Classification versus the Angular Displacements for the Four-Class Problem.*

In conclusion, the M-P classification performance has been described for two and three perspectives and applied to a three- and a four-class problem. The multiple perspectives affect the single class probability of correct classification differently depending mainly on the number, nature and persistency of scattering centers appearing in the profiles. As a consequence, the node location dependence of the global classification rate decreases when a greater number of classes is involved, making the classifier equally reliable for a wide range of angular displacements. The M-P classification improvements are reduced when the nodes are closely separated since the perspectives exhibit a significant degree of correlation. Nevertheless, the overall probability of correct classification is well above the mono-perspective case (+11.3%

and 15.6% using two and three perspectives respectively). In addition the complexity and variability of reflectivity from real targets has been highlighted. Multiple perspective classification does not necessarily offer a trouble free route to acceptable classification and requires further testing under more realistic conditions. It also helps to indicate what information in the radar signature is important for classification. However, much further research remains before routine and reliable classification by radar becomes the norm.

5.4 MSTAR data: Aspect Dependent Drivers for ATR

The effects of multiple perspectives are here further analysed on test invariant targets and confusing classes as discussed in Section 4.2 in order to describe more accurately the performance improvements using those parameters introduced in Chapter 4 (generalisation and ROC curves analysis). Therefore, the MSTAR set of two-dimensional SAR data (see Section 4.3.1) has been tested using multiple perspectives.

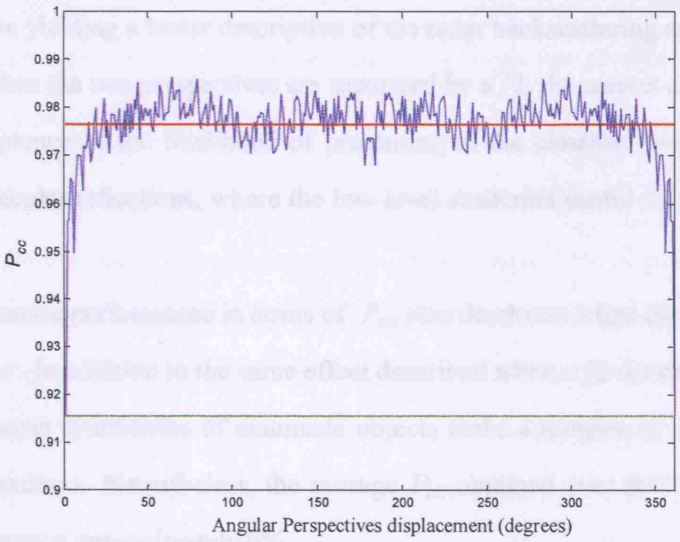

Figure 5.17: *Probability of correct classification of a two-perspective classifier at different angular displacement $\Delta\varphi_{1,2}$ in a forced decision environment.*

Again, although the target has already been assumed to have been detected and tracked, the angular displacements of the nodes are unknown to the hypothetical network collecting

the perspectives. The network topology Φ, i.e. the angular aperture between the nodes, is not processed as information by the Multi-Perspective classifier. Whilst this is only one of the possible approaches to multiple views of the target it is relatively simple to implement and enables the performance against multiple perspectives to be investigated.

Furthermore, the computational burden is reduced to the minimum. In figure 5.17, the correct classification rates are plotted versus the angular displacement of the perspective for a forced decision environment. The green line is the result of a two-perspective classifier with zero degrees separation (i.e. a single-perspective case). The two-perspective (blue) P_{cc} emphasises that as soon as the two perspectives decorrelate, the information content of the two signatures increases with the result of aiding the classification task. The average over all 2-P topologies (red) demonstrates the efficiency of the second perspective contribution compared to the single-perspective case (green). The probability of correct classification measured by the single-perspective classifier can be observed when the two sensors are co-located (i.e. when $\Delta\varphi_{1,2} = 0°$). As observed for the ADAS data, as $\Delta\varphi_{1,2}$ increases, the two perspectives start to decorrelate yielding a better description of the radar backscattering and a performance improvement. When the two perspectives are separated by $\pi/2$, the correct classification rates drop as a consequence of the likelihood of presenting to the classifier two signatures both dominated by specular reflections, where the low level scatterers useful for classification are obscured.

The classification performance in terms of P_{cc} also decreases when the two perspectives are separated by π. In addition to the same effect described when $\pi/2$ occurs, in this case the number of geometric symmetries of manmade objects make a number of signatures similar at π separated headings. Nevertheless, the average P_{cc} obtained over 360° (red) is 97.65%, showing an increase of approximately 6%.

In Fig. 5.18, the graphs related to the ROC curves are shown. In particular, the mean value (red) of all the possible angular displacements between the two radar locations over 360 degrees can be compared to the single-perspective case (green) and other perspective location performance (5°, 20°, 55°).

The ROC curves show an overall improvement with respect to the single-perspective

case although a slight dependence on the location of the nodes. As regards the ROC curve "b", the benefits of perspective diversification are valuable also for small angular displacements between the perspectives, while the multi-aspect effectiveness on the relationship between and requires a greater perspective separation.

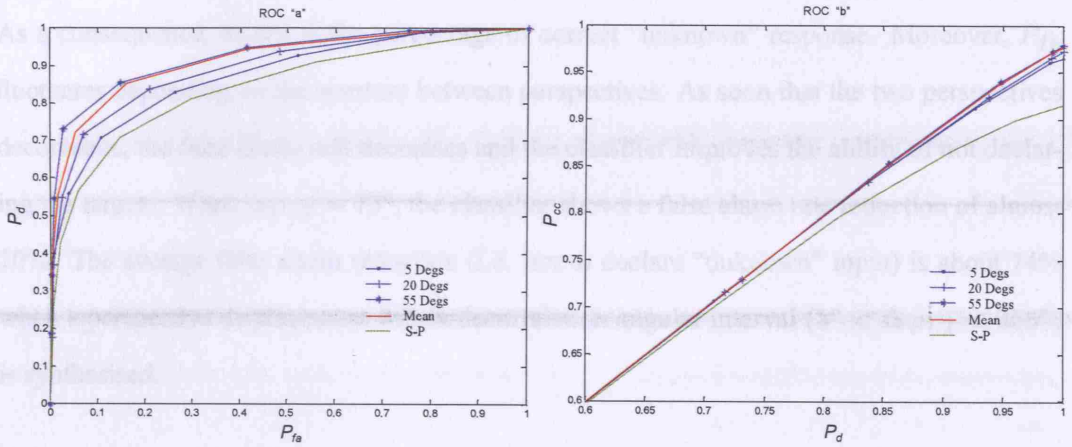

Figure 5.18: *ROC curves "a" and "b" for a single-perspective classifier (green), mean value over 360° (red) achieved with two perspectives and two-perspective performance using different angular displacements (5°, 20°, 55°).*

The reliability, which can be measured from the slope of ROC curve "b" (see Section 4.3.3) also significantly increases using a further perspective. Further perspectives aid the classifier to not declare the unknown targets, therefore reducing the false alarm rate at a fixed probability of declaration. Moreover, this is translated in ROC curve "b" as a higher degree of confidence when a declaration is made. Interestingly, it might seem that the second perspective benefit on classification performance is less dependent on the angular displacement between perspectives if compared to the AUC of ROC curve "a". To understand better the classification improvements, in fact, the two curves need to be related to each other: after fixing the acceptable false alarm rate, the corresponding probability of declaration is then used to find the correct classification rate related to that parameter. For example, if a $P_{fa} = 0.6$ is fixed, the single-perspective classifier guarantees that $P_d = 0.92$ and, from the curve "b", a $P_{cc} = 0.87$. By using a second perspective, the probability of declaration becomes $P_d = 0.97$

leading to a $P_{cc} = 0.96$, showing an increase of correct classification of about 9%.

In order to investigate the false alarm reduction on those targets not depicted in the training set, the classification performance on the "unknown" targets is now investigated. In figure 5.19, the improvement $(1-P_{fa})$ of not declaring the confusing targets D7 and BTR60 is shown for a particular rejection threshold leading to a $P_{fa} = 55.2$ % in the single perspective case. As a consequence, 44.8% is the percentage of correct "unknown" response. Moreover, P_{fa} fluctuates depending on the aperture between perspectives. As soon that the two perspectives decorrelate, the false alarm rate decreases and the classifier improves the ability of not declaring the targets. When $\Delta\varphi_{1,2} = 75°$, the classifier shows a false alarm rate reduction of almost 20%. The average false alarm reduction (i.e. not to declare "unknown" input) is about 14% when a perspective displacement within decorrelation angular interval $(5° < \Delta\varphi_{1,2} < 355°)$ is synthesised.

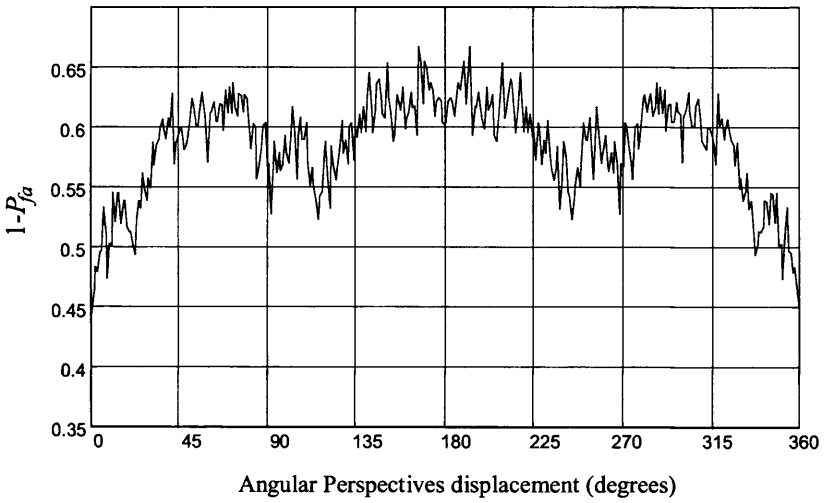

Figure 5.19: *Probability of not declaring unknown targets* $(1 - P_{fa})$ *as a function of the perspective displacement* $\Delta\varphi_{1,2}$.

The Multi-Perspective benefits on the generalisation capabilities of the classifier can be deduced for the test variant targets ($T72_{ind}$ and $BMP2_{ind}$). In figure 5.20, the probability of correct classification in a forced decision implementation is shown for the test and test variant targets. The mean value of probability of correct classification over the perspective displace-

ment is 96.7%, that means 6.1 % improvement when compared with the single perspective (1-P) case.

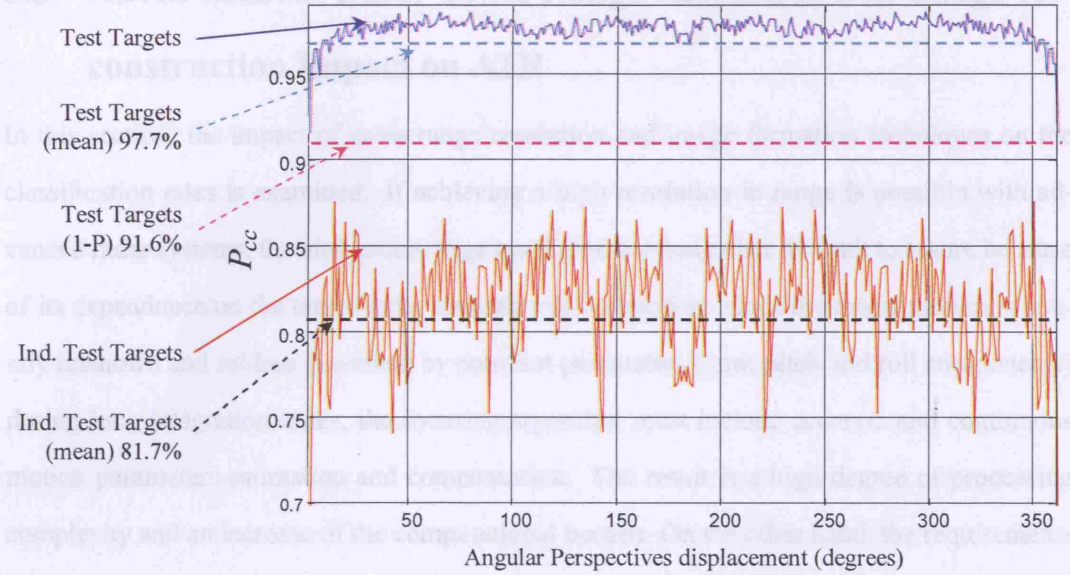

Figure 5.20: *Generalisation and correct classification improvements given by different perspective displacements.*

The generalisation capabilities show a more pronounced increase (+11.8%), from 69.9% to 81.7% overall, which can reach +15% for some favourable perspective displacements. The drop of correct classification coincides with a drop of generalisation capabilities in the neighbourhood of $\Delta\varphi_{1,2} = \pi$. Once again, as described for the ADAS-MIDAS analysis, this is due to the symmetries of manmade targets and the corresponding likelihood of observing two highly correlated signatures when this perspective topology is formed.

In conclusion, the results of employing multiple signatures sensing a radar target have been presented using 2-D images of a six-class population set. The ROC curves have been described, showing the considerable benefits offered by multiple perspectives. Furthermore the perspective diversification effectiveness on the relationship between false alarm rate and probability of declaration requires a greater perspective separation if compared to the one between declaration and correct classification. The independent targets also show a valuable

increase in correct classification rates if compared with the results obtained using a traditional single view of the target.

5.5 ADAS-MIDAS Data: Cross Range Resolution and Image Reconstruction Impact on ATR

In this section, the impact of cross range resolution and image formation techniques on the classification rates is examined. If achieving a high resolution in range is possible with advanced radar systems, the high cross-range resolution is often rather difficult to assure because of its dependence on the target-radar antenna relative motion: since the target motion is usually unknown and seldom described by constant parameters (yaw, pitch and roll components) during long integration times, the focusing algorithm must include accurate and continuous motion parameters estimation and compensation. The result is a high degree of processing complexity and an increase of the computational burden. On the other hand, the requirements on cross-range resolution in order to separate target scatterers lying in the same range bin could be less demanding than in the slant-range dimension. As a consequence of this asymmetry of cross and slant range resolution needs for target recognition (Rihaczek and Herschowitz, 2000), the question concerns the best way of combining the 1-D signatures in case of large synthesised apertures.

Three different processing techniques are examined from a classification perspective. The first method involves a focusing algorithm on the entire sequence of N profiles. The second technique, after dividing the N profiles into different sets representing the sub-apertures, consists of giving multiple signatures corresponding to different looks as input to the classifier. The last approach differs from the second as the final pattern to be recognised is the multi-look reconstruction (non-coherent superposition of rotated images) of the N sub-apertures. In this last case the cross range resolution starts losing its definition since the progressive rotation of the cross range direction and its combination with the slant range. The information content of the patterns used for training and testing the classifier is derived from the same number N of collected radar echoes. The first and last approaches require the target motion to be accurately known while the multi-signature processing necessitates the target motion information

for scaling purposes only. As for the first method, the main drawback of the Multi-Look reconstruction approach lies in the computational burden. Conversely, the second method can be performed with simple DFT processing on small apertures and is particularly suited when simplicity and high computational velocity are required. The classifier, on the other hand, needs to be tailored to accept more than one image as input and its performance is strictly connected to the changeability of the target signature in the aperture integration time.

The data used for this purpose are the ADAS-MIDAS set of six targets (Section 4.4.1). The range resolution is $\sim$ 8 cm and, after forming the 2-D images using the three different techniques mentioned above, PCA is applied. As described in the previous sections, the information content of ground vehicles radar signatures displaced by 180 degrees on the same projection plane shows a minimum from a classification perspective. As a results, 11 sets of profiles over 360°, one set every 32°, have been selected, characterising the radar backscattering over apertures $\Psi = 12°$. The test set consists of signatures covering the same aperture Ψ, displaced by one degree, neglecting the ones used to train the classifier.

5.5.1 Coherent image focusing technique

The full resolution radar target images are obtained by coherently processing the radar echoes over an aperture of about $\Psi = 12°$. Since, at the considered resolution, the scatterer migration through range and Doppler cells is not negligible a proper algorithm has to be applied to cope with it. In this work, to focus the images of the considered ground targets, we used an approach as in Aprile et al. (2004) properly modified to take into account for the turntable specific conditions. The main steps of the focusing technique are here briefly summarised: (i) first compensate for the range migration dependent on the cross-range position of the scatterer with respect to, w.r.t., the rotation centre, (ii) then compensate for the Doppler migration dependent on the slant range displacement w.r.t. the rotation centre, and (iii) finally compensate for the Doppler migration dependent on the cross-range displacement w.r.t. the rotation centre. The rotation centre coincides with turntable centre fixed w.r.t. the radar antenna. It could be shown that step (iii) can be neglected when the rotation rate is sufficiently constant during the time aperture. An example of full resolution radar image is shown in figure 5.21.c as the input of

the coherent image classifier. By comparing this image with the corresponding 20 cm cross-range resolution images shown in the same figure as inputs to the multi-signature classifier, it can be noted that the coherent processing allows the different scattering centres to be properly resolved thus providing high quality images to the subsequent classification procedures.

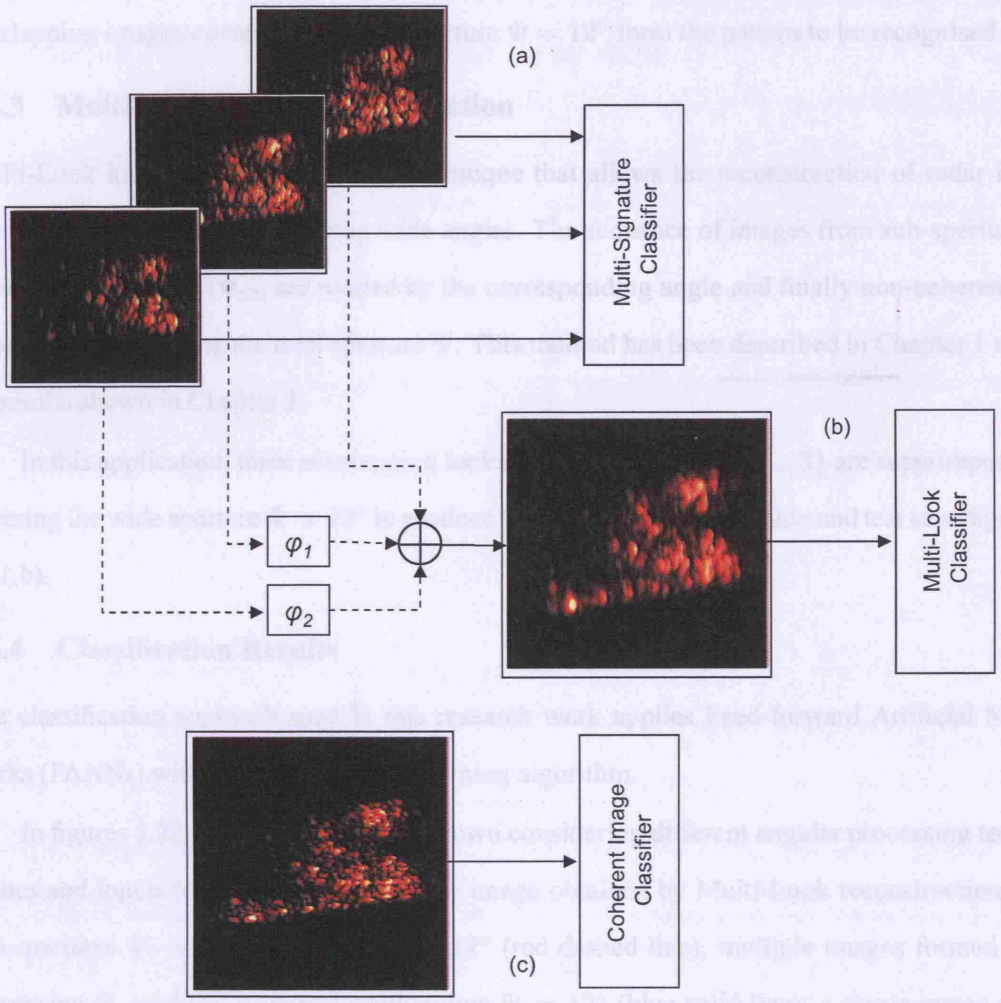

Figure 5.21: *Three angular processing techniques used to test the classifier (Multiple signatures, Multi-Look image reconstruction and Coherent image focusing technique).*

5.5.2 Multiple Signatures

A limit of 2-D DFT to perform complete and fully focused target images is the migration through different cells: when the rotation angle becomes greater than a certain limit, a point

scatterer changes range bin, its Doppler information deteriorates and the scatterer's energy falls into many different range/Doppler cells causing a blurred image. The range bin migration can be overcome by processing data collected from small arcs of circular aperture Ψ_i, where also the Doppler shift can be assumed constant. Therefore, a series of images covering the aperture Ψ is given as input to the classifier as illustrated in figure 5.21.a. Specifically, three non-overlapping images covering the large aperture $\Psi = 12°$ form the pattern to be recognised.

5.5.3 Multi-Look Image Reconstruction

Multi-Look image reconstruction is a technique that allows the reconstruction of radar images of rotating targets synthesising wide angles. The sequence of images from sub-apertures previously described (Ψ_i), are rotated by the corresponding angle and finally non-coherently superimposed covering the total aperture Ψ. This method has been described in Chapter 1 and its results shown in Chapter 3.

In this application, three overlapping looks $\{\Psi_i = 4.25° : i = 1, ..., 3\}$ are superimposed covering the wide aperture $\Psi = 12°$ to produce the elements of the training and test sets (figure 5.21.b).

5.5.4 Classification Results

The classification approach used in this research work applies Feed-forward Artificial Networks (FANNs) with back-propagation learning algorithm.

In figures 5.22, the ROC curves are shown considering different angular processing techniques and inputs to the classifier: a single image obtained by Multi-Look reconstruction of sub-apertures $\Psi_i = 4.25°$ covering $\Psi = 12°$ (red dashed line); multiple images formed by processing $\Psi_i = 4.25°$ each and synthesising $\Psi = 12°$ (blue solid line); a single image obtained by coherently processing the entire aperture $\Psi = 12°$ yielding $\Delta r_c = 6.67$ cm reduced to ~ 8 cm after weighting (green dashed-dotted line). The wide aperture performance can be collated with the classifier using a single image formed from an aperture $\Psi_i = 4.25°$ (black dotted line).

If the probability of false alarm is 0.4, for example, different P_d values are obtained for the different techniques, yielding $P_{cc} = 74.6\%$ for the small aperture technique, 78.4% for the

Multi-Look, 78.7% for the multi-signature and 87% for the coherent processing.

The effectiveness of the latter technique can be demonstrated for other probabilities of false alarm. This can be deduced from the AUCs of the different techniques: although comparable for the ROC curves "b", the coherent processing exhibits a noticeably larger area of the ROC curve "a". On the other hand, the coherent image formation requires more principal components to describe the information content of the image.

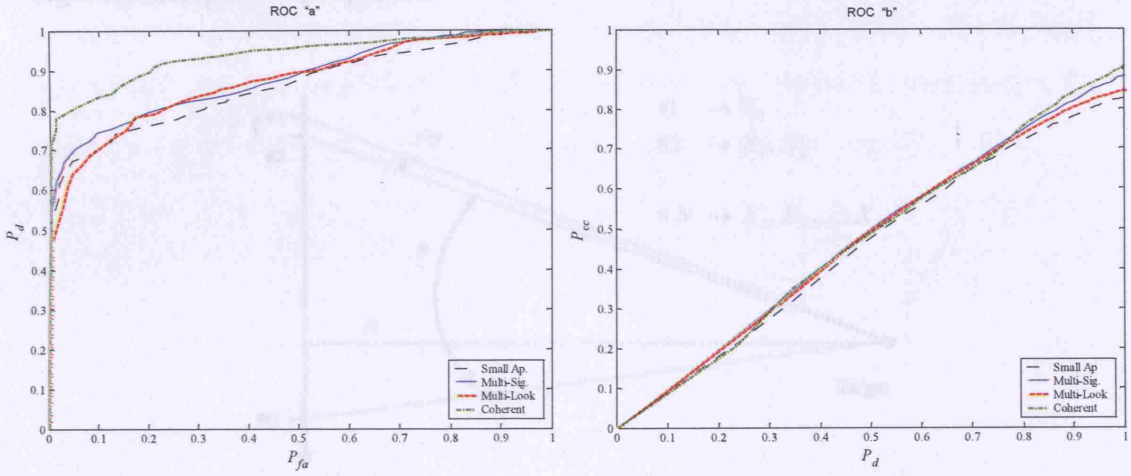

Figure 5.22: *ROC curves for the three angular processing techniques.*

The multi-signature technique shows a slightly higher classification response than the Multi-Look processing, suggesting that the noise and speckle removals balance the merging of the cross and slant range resolutions and the averaging of those features appearing at different observation angles in a single image.

The classification performance of the techniques using the backscattering information from large apertures is well above the small apertures case for any rejection threshold fixed.

The classification results demonstrate the reliability augment when the coherency is maintained in reconstructing the image. Nevertheless, ROC curves shown by other techniques indicate the convenience of processing consecutive looks over the single sub-aperture case.

5.6 Application

We now examine a possible application of the Multi-Perspective classification approach, which can be identified with the multiple signatures collected by an airborne system flying past a target. If a circular trajectory is covered around the object of interest, the ISAR geometry described for the ADAS-MIDAS data can be thought of as an approximation of the problem. In a real scenario, the non-circular trajectory would change the distance to the target and therefore the SNR with the consequence of having qualitatively different descriptions of the target depending on the perspective.

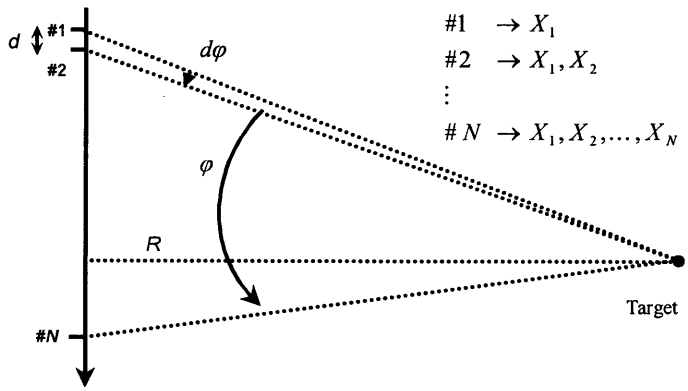

Figure 5.23: *Cumulative acquisition of perspectives. The system collects N signatures spaced by a distance d and a variable perspective displacement dφ.*

On the other hand, an arbitrary trajectory would change the depression angle with the result of decorrelating the perspectives more than at constant elevation, yielding more pronounced Multi-Perspective information enhancements and therefore performance improvements. This aspect might also affect the construction of the library set, which must be expanded in order to represent the target from all possible three-dimensional orientations.

We assume an aircraft flying a linear trajectory as in figure 5.23, at a velocity of 200 m/s at a range $R = 10$ km from the target and transmitting $n = 8$ chirps using stepped-frequency compression techniques at PRF= 3 kHz. Considering the ADAS system described in Chapter 3, the number of collected signatures at the same centre frequency and polarisation, after

averaging over 4 profiles would be $N \sim 45$ per second, that is one profile every $d = 4.4$ m.

A Multi-Perspective classifier can be implemented on a *cumulative* basis, i.e. the HRR profiles $X_1, X_2, ..., X_N$ are continuously collected and progressively presented to the classifier, or on an *interrupted* basis, i.e. the profiles are taken from ad hoc perspectives, chosen according to the best probability of correct classification suggested by the Multi-Perspective graphs (figures 5.15 and 5.16). In the interrupted mode of operation, between consecutive collections the system can be switched off, reducing the probability of being detected by passive sensors. In figure 5.24, the classification performance of the cumulative classifier is shown for an aperture $\varphi = 60$ degrees (i.e. the radar is "on" for almost one minute). After the rapid increase of the correct classification rate due to the progressive signatures decorrelation, the updated score increases linearly with the synthesised aperture ($\sim 5\%$ every $25° - 30°$).

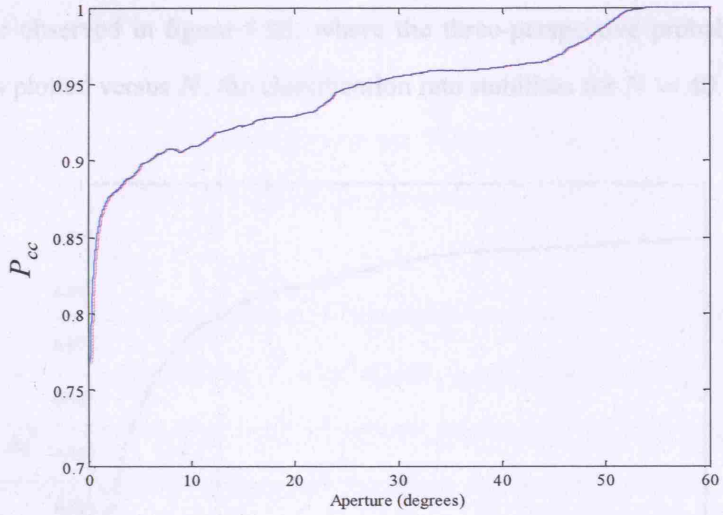

Figure 5.24: *Cumulative M-P correct classification rates versus covered aperture.*

The interrupted geometry is depicted in figure 5.25. The radar system is "on" for a limited amount of time so that N profiles are collected. Then the system is switched off until the next perspective is reached. The perspective displacements vector $\Phi = \{\Delta\varphi_{1,2}; \Delta\varphi_{1,3}\}$ is chosen on the basis of the best P_{cc} obtained for a three perspective classifier for a total aperture of 60°. This is verified when $\Phi = \{28°; 60°\}$, and $P_{cc} = 91.6\%$ (see figure 5.16).

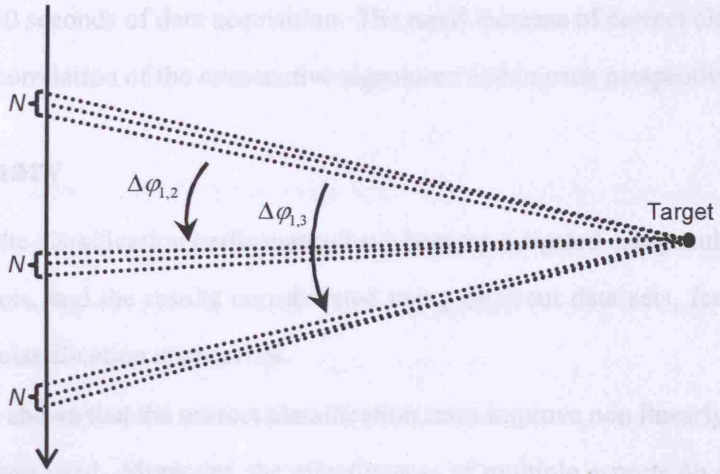

Figure 5.25: *Interrupted acquisition for M-P classification. The system can be switched off between consecutive perspectives, reducing the probability of being detected.*

As can be observed in figure 5.26, where the three-perspective probability of correct classification is plotted versus N, the classification rate stabilises for $N = 40$.

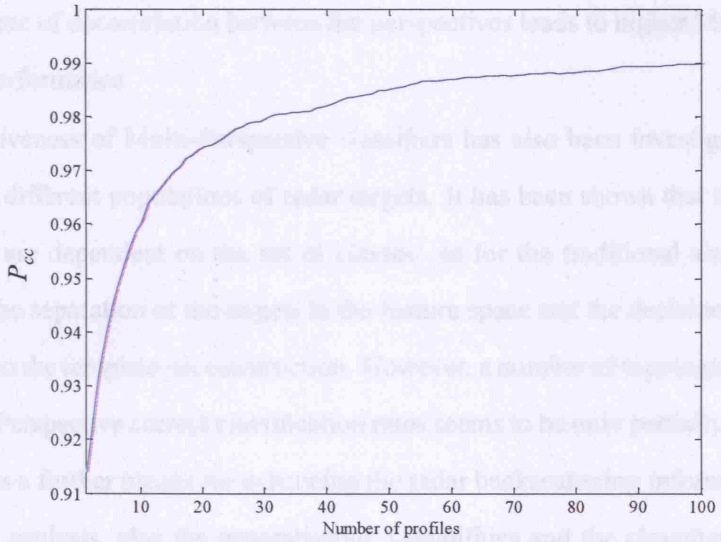

Figure 5.26: *Interrupted M-P P_{cc} versus number of signatures per perspective.*

This is achieved when the radar is "on" for a time $t \sim 1$ s for each perspective. In this case the classifier shows a probability of correct classification comparable to the cumulative

classifier after 40 seconds of data acquisition. The rapid increase of correct classification rate is due to the decorrelation of the consecutive signatures within each perspective.

5.7 Summary

In this chapter, the classification performance have been investigated using multiple signatures from radar targets, and the results corroborated using different data sets, feature extraction algorithms and classification approaches.

It has been shown that the correct classification rates improve non linearly with the number of perspectives used. Moreover, the effectiveness of multiple aspects on classification is significant also for low SNR, recovering a portion of the low level scatterers information lost with the noise level increase.

The role of the perspective displacement and the target intrinsic symmetries and characteristics have been also introduced through the examination of the probability of correct classification in 2-P and 3-P topologies. The cross-correlation between target signatures at different orientation aspects also provide an insight into the backscattering information content. Usually, a high degree of decorrelation between the perspectives leads to higher Multi-Perspective classification performance.

The effectiveness of Multi-Perspective classifiers has also been investigated testing the classifier using different populations of radar targets. It has been shown that the correct classification rates are dependent on the set of classes: as for the traditional single-perspective classification, the separation of the targets in the feature space and the decision boundaries are strictly related to the template set construction. However, a number of topologies yielding high and low Multi-Perspective correct classification rates seems to be only partially influenced and could be used as a further means for enhancing the radar backscattering information. Through the ROC curve analysis, also the generalisation capabilities and the classifier reliability improve using multiple views of the target.

In case of contiguous perspectives collected by a single sensor, different techniques of radar image processing have been discussed showing that the full coherent processing guarantees the most detailed backscattering description. Nevertheless, the Multi-Look and Multi-

Signature introduced show robust classification results. These techniques are more feasible to be implemented in real systems since they require a lower computational burden and processing complexity.

Finally, two possible applications examined. The multiple perspectives can be *ad hoc* selected on an interrupted basis or continuously collected by the system on a cumulative basis. The algorithms developed can therefore be incorporated into existing systems.

Chapter 6

The Knowledge of the Perspective Displacement

In the previous chapter, the knowledge of the perspective disposition Ψ (i.e. the angular displacement between looks at the target) has been assumed unknown to the Multi-Perspective system. This scenario is likely to happen when the target is moving inside the radar field and changing its orientation during the time-on-target. If the target is not tracked and its motion parameters cannot be estimated, a moving system collecting radar signatures from different perspectives at different times is not able to estimate the M-P topology accurately.

Multi-Perspective (M-P) radar ATR can be implemented following different approaches. As described in the previous chapters, after detection, clutter discrimination and segmentation, a sequence of signatures representing different views, either 1-D or 2-D imagery, can be thought of as a set of perspectives $X_1, X_2, ..., X_N$. M-P *disjoint* classification can be implemented using this information only. The required pre-processing does not affect the traditional classification process since it only involves a series of single-perspective classifiers independently producing a partial score, which is finally combined in order to make a decision (declare or not-declare).

If the system is able to compute the perspective displacement of the perspectives in azimuth and elevation then the M-P *joint* classification can be considered. A number of target positioning techniques have been developed, such as Time Of Arrival (TOA, also known as *trilateration*), Time Difference Of Arrival (TDOA, also known as *multilateration*), Angle Of

Arrival (AOA, also known as *triangulation*) and hybrid methods processing time and bearing information. These techniques can be implemented using netted radars or individual systems performing *lobing* or *monopulse* operations (Stimson, 1998). The former technique involves sequential pulses alternating the mainlobe direction so that the angle measurement of the target is obtained by comparing the intensity of the received echoes. Differently, monopulse systems estimate the target angle of arrival using a single pulse and either producing simultaneous lobes (amplitude comparison) or a single lobe using two spaced antennas in receive (phase comparison).

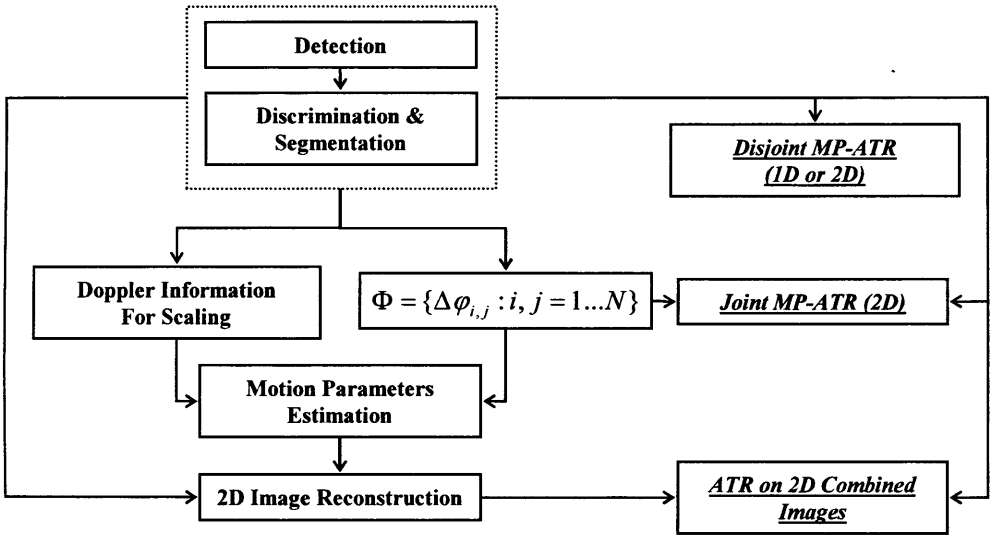

Figure 6.1: *M-P classification approaches.*

Finally, the possibility of connecting the perspectives, together with the system/target motion estimate, allows signature reconstruction and the formation of a combined image for ATR (figure 6.1). Non-coherent image recombination over uncorrelated perspectives has the advantage of reducing noise and speckle, but it is effective for point-like features only (i.e. optical images) since in radar images it might merge the main radar scattering centres radiating pattern information. For example, if a target exhibits a corner-like scatterer when sensed at a particular heading, this potentially discriminative feature is averaged by different perspectives and, when combined in a single image, would loose the link to the particular orientation yield-

ing that response. Furthermore, in case of projection plane change (i.e. the perspectives taken at different azimuth and elevation angles) the resulting 3-D image would require a significant processing complexity.

Therefore, the discussion is here focused on the other possible techniques (figure 6.1) based on individual perspectives: *disjoint* M-P classification (investigated in Chapter 5) and its *joint* implementation which uses the perspective disposition as further information.

6.1 Significance of the Angular Perspective Displacement

We now concentrate on the scenario where the M-P system processes the perspective displacement between multiple views. For example, a network of radars collecting multiple signatures at the same time from different perspectives, while the target can be assumed stationary. Another application can be found in moving radars sensing a tracked target over time, whose motion parameters can be estimated. It is clear that the estimation of the platform and target motion parameters increases the system complexity and requires further processing, although this is counterbalanced by the extra information gained, whose benefits are investigated here.

As previously stated, the image recombination from multiple perspectives would lead to reradiating pattern information loss from individual scattering centres. Furthermore, it would require a high complexity due to the scaling and registering processing. Therefore, a possible approach would be to assign a scattering state that identifies a particular range of target headings. Hence, the perspective displacement can be thought of as an indicator for the state transition. This approach has been developed using Hidden Markov Model (HMM), which is described in detail in Appendix D.

Each class C_i is identified by a model $\lambda^{C_i}(A, B, \pi)$ Specifically, as stated by (Shihao et al., 2005), if a state S_i is defined for a range of aspect angles $\alpha_i = [\alpha_{i,1}, \alpha_{i,2}]$, and assuming that a variable aperture $\widehat{\alpha}_i = \alpha_{i,2} - \alpha_{i,1}$ identifies different states, the initial state probability π_i can be calculated for each class as follows:

$$\pi_i = \frac{\widehat{\alpha}_i}{\sum_j \widehat{\alpha}_j} \tag{6.1}$$

Initially, (Runkle et al., 1999b), the HMM was modelled as discrete, i.e. the input vector

was quantised and associated with a symbol in a codebook. Moreover, the perspectives were assumed to be contiguous and therefore the sequence of states was correlated. However, here we are interested in investigating the classification performance when the perspectives exhibit a high degree of decorrelation. We also model the observation as continuous.

As described by (Shihao et al., 2005), the transition probability between the i-th and j-th states is regulated by the angular displacement between the looks:

$$a_{i,j}(\varphi) = w_j\left(d_{i,j} - \varphi\right) \tag{6.2}$$

where φ is the measured angular separation between two perspectives. The weighting function w_j is maximum when the measured perspective displacement is equal to the angular distance $d_{i,j}$ between the two states:

$$w_j\left(\beta\right) = \frac{1}{\sqrt{2\pi(\sigma_j)^2}} \exp\left[-\frac{1}{2}\left(\frac{\beta}{\sigma_j}\right)^2\right] \tag{6.3}$$

where the parameter σ_j is chosen to maximise the final probability of correct classification of the sequence.

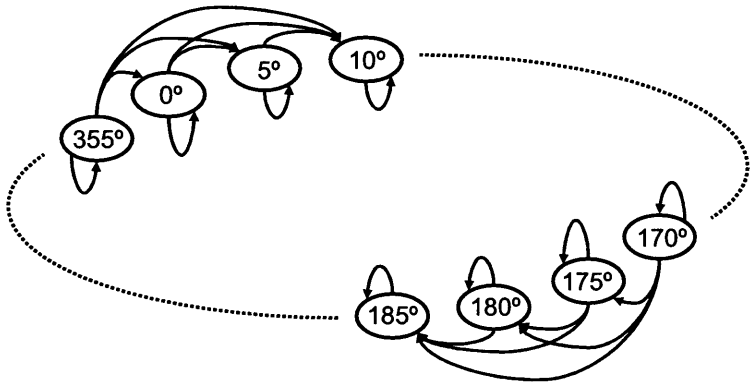

Figure 6.2: *"Left-Right" HMM structure. Each state S_i represents the backscattering of the target over 5 degrees in azimuth. The elevation dimension is not considered.*

We start our description adopting the same transition and initial state models. We also assume the same range of aspect angles identified by different states: $\widehat{\alpha}_i = \widehat{\alpha},\ 1 \leq i \leq N$ where N is the number of states in which the backscattering is subdivided and represented.

In figure 6.3 a "Left-Right" HMM implementation of joint M-P classification is shown. The target orientation covered by each state is $\widehat{\alpha} = 5°$ and the structure is circular: when the angular perspective is measured, the transition matrix A is evaluated according to equation 6.2 and 6.3.

In order to characterise the model, the state distributions B consisting of the probabilities $P(X|S_i)$ related to the states $i = 1, 2, .., N$ need to be defined (figure 6.3). A number of possible approaches can be followed. When the emission is modelled following the approach suggested by Shihao et al. (2005), i.e. using a Gaussian model, the classification performance achieved using the ADAS-MIDAS data and the template library and test set introduced in Section 5.3 are compared here using the perspective information $\Delta\varphi_{1,2}$. In order to meet the same library information characteristics for both joint and disjoint M-P classification, 36 states are represented, one every 10 degrees.

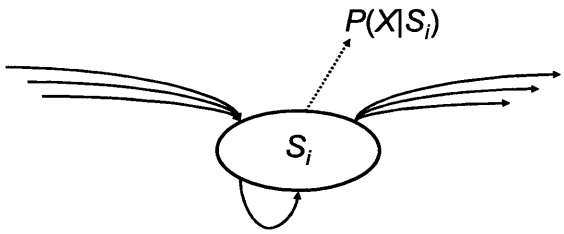

Figure 6.3: *Emission for a given HMM State. The probability of generating a vector (symbol) X in a state S_i can be modelled in different ways.*

The state representation is strictly related to the library of templates. The correct classification rate performance is shown in figure 6.4. The overall correct classification rates of the joint classifier using HMM with the Gaussian state distribution are less than the disjoint approach using the approach discussed in Chapter 5. This is due to the state representation and the corresponding emission probability being based only on a single template vector. When the disjoint M-P approach is used, the target class information resides in the full 360 degrees information and it is possible that some of the features have a persistency over more than a state, or being occluded for a few states for appearing again in other states. Their characterisation can be therefore generalised more accurately when the full set of templates is used instead

of combining the emissions of individual states. As a consequence, the disjoint classification performance in terms of probability of correct classification is higher for a target class when a test target heading is between two template orientations.

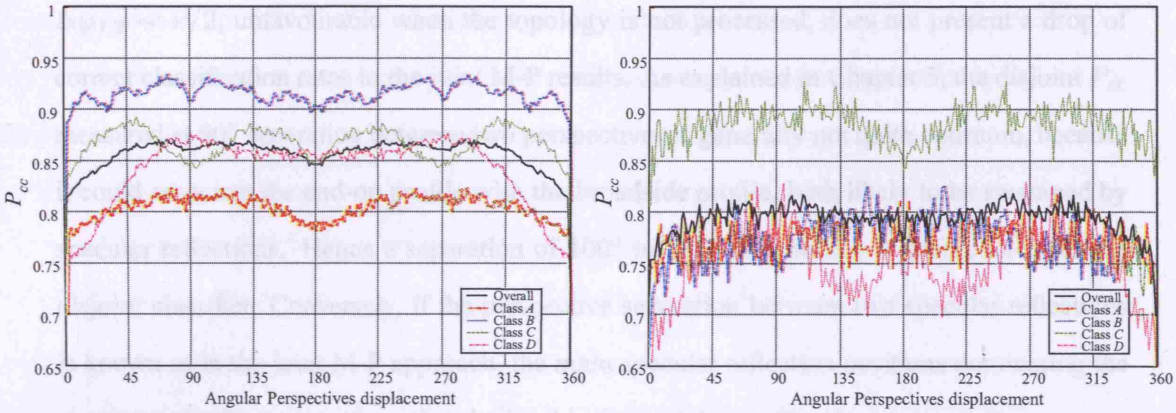

Figure 6.4: *"Disjoint" (left-hand side) and "Joint" (right-hand side) Classification Comparison. The knowledge of the perspective topology is exploited using HMM.*

The consequence of the reduced information with state representation are also shown by the fluctuations of the single class P_{cc}. The closer the test vector to the centre of the state, the higher the discrimination performance. As a conclusion, the HMM based M-P classification requires a higher information content in order to achieve the same performance as the disjoint M-P classification. Therefore, each target needs to be represented by more than one template in each state when HMM classification is used in order to compute a more reliable estimate of the emission probabilities.

The P_{cc} trend with the perspective displacement for the different targets illustrates similarities between the results from the two different Multi-Perspective approaches. Firstly, the decorrelation requirements are similar for the overall performance (black solid line) and a few degrees of perspective aperture are necessary in order to gain the multi aspect benefits. The angular decorrelation exhibited by the class D target is greater than the other classes as already observed for the disjoint classifier. Furthermore, $\Delta\varphi_{1,2} \sim \pi$ shows a drop of classification performance as a consequence of the symmetries of ground-vehicles at this particular perspective

separation.

The differences of the results from the different methodologies are also significant. Although the state fluctuations, the joint M-P approach shows less overall variability with respect to the particular perspective displacement. For instance, the angular displacement $\Delta\varphi_{1,2} \sim \pi/2$, unfavourable when the topology is not processed, does not present a drop of correct classification rates in the joint M-P results. As explained in Chapter 5, the disjoint P_{cc} measured at 90° separation between two perspectives is generally not quite optimum, because it could associate the end-on profile with the broadside profile, both likely to be swamped by specular reflections. Hence a separation of 100° to 110° is generally a good choice for the disjoint classifier. Conversely, if the perspective separation between two specular reflections is known as in the joint M-P approach, the main specular reflection positions dominating the signatures for the target orientation depicted in figure 6.5 can either be interpreted as:

- **6.5.a**: Location refinement of the specular reflection range of the same broadside or end-on target heading with respect to the centre of target rotation $\Phi = \varphi_{1,2} = 2k\pi,\ k = 1, 2, ...$

- **6.5.b**: Shape measurement given by one broadside and one end-on specular reflection when $\Phi = \varphi_{1,2} = \pi/2 + k\pi,\ k = 1, 2, ...$

- **6.5.c**: Target radar length estimate along one of the two axes of symmetry when $\Phi = \varphi_{1,2} = \pi + 2k\pi,\ k = 1, 2, ...$ In this case either two broadside or two end-on measurements are processed.

In figure 6.5 it has been assumed that the end-region scattering mechanism at the far end of the target is occluded and therefore a single scattering centre dominates the signature, as commonly verified for low grazing illuminating angles.

For this particular scenario, it is clear that the knowledge of the perspective displacement can be a significant aid to the M-P classification since the information on the shape of the target can be more accurately described since the cases (a) and (c) are equivalent for the disjoint approach, while distinguished for the joint implementation. Nevertheless, the backscattering

of each state S_i requires accurate representation in order to achieve sufficiently high emissions $P(X|S_i)$ and therefore classification performance. This approach is suitable when there are no restrictions to the number of templates that can be used to depict the class states. On the other hand, the disjoint approach exhibits more robust classification performance when only a few patterns are used to represent the target class since the higher capabilities of generalising the backscattering features between the headings represented in the template library.

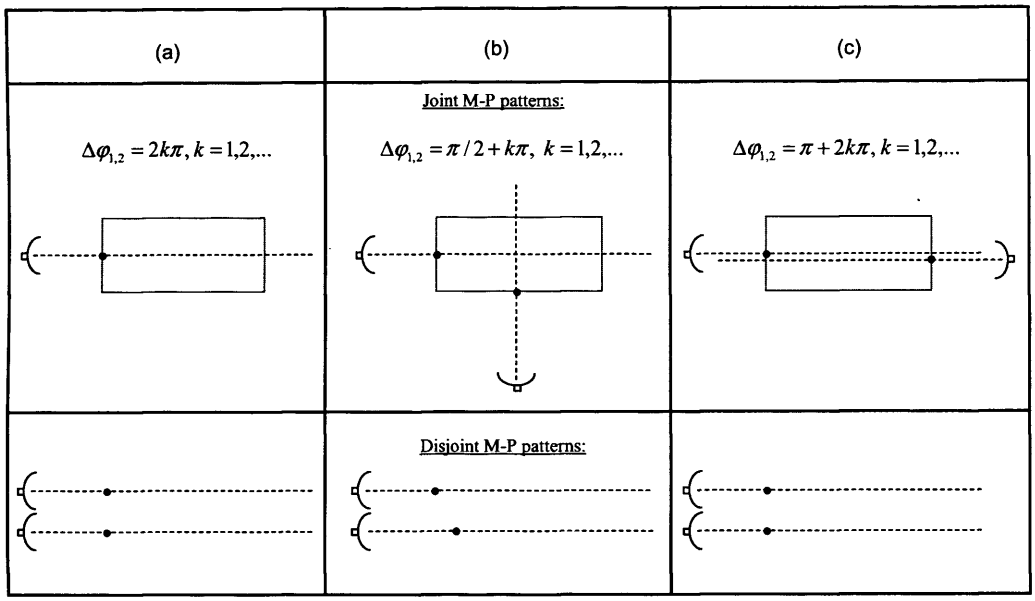

Figure 6.5: *Association of HRR profiles using the perspective topology information ("joint M-P" classification) and the only profiles ("disjoint M-P" classification).*

6.1.1 Multi-Perspective Classifier using Neural Networks: Hybrid HMM

As previously described, continuous HMM can be used in Multi-Perspective classification, although care must be taken both on the emission probability estimate and the template set used to represent the state decomposition of target orientations.

Artificial Neural Networks (ANN) can be used in order to compute the state probabilities $P(X|S_i)$ and combined with HMM classification. This Hybrid Hidden Markov Model approach has been used in speech recognition (Bourlard and Morgan, 1994) and is discussed later on.

As regards the target state representation, HMM requires a detailed state description in order to accurately estimate the likelihood of the signature X to be generated in state S_i. Hence, the MSTAR set of data (see Chapter 4) is used for this purpose since the training set is formed by an exhaustive set of templates, almost one SAR image every degree of target heading. Therefore, $\sim$ 5 templates from each state can be used to represent a state. Unfortunately, some of the states have no template representation and it is therefore necessary to expand the state aperture to 20 degrees.

The *forward-backward* algorithm described in Appendix D is used to evaluate the probability $P(X_1, ..., X_N|\lambda_i)$ that the sequence $X_1, ..., X_N$ has been generated by the model λ_i representing the class C_i.

We now calculate the emission probabilities $P(X_i|S_i)$ using the neural network. The architecture of the hybrid HMM consists of a number of states identified by individual neural networks and trained on the templates presenting the orientation in the range of angles that the state represents (figure 6.6). The i-th state network S_i is trained on the template set from each class within the heading interval $[\alpha_i(1), \alpha_i(2)]$:

$$\mathrm{T}(S_i) = \bigcup_{k=1}^{n_c} T(S_i, k) \quad \text{where } T(S_i, k) = \left\{ \mathbf{t}^k_{\alpha_i(1)}, ..., \mathbf{t}^k_{\alpha_i(2)} \right\} \tag{6.4}$$

The posterior probability $P(S_i|X)$ is estimated as the output of the FANNs and inverted into likelihood using the Bayes formula.

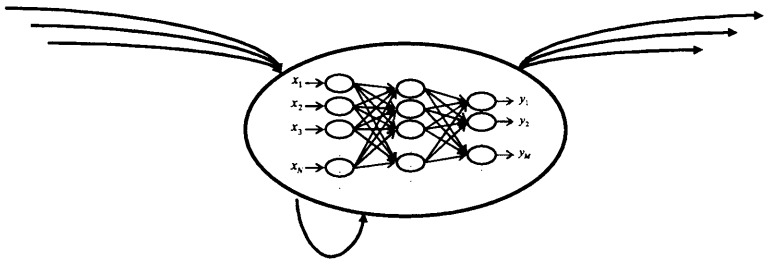

Figure 6.6: *Hybrid HMM/ANN model for the state emission probability.*

The initial state probabilities π_i are calculated using equation 6.1, while the transition probability matrix $A = a_{i,j}$ is evaluated using equations 6.2 and 6.3.

The results of the Hybrid HMM/AMM based on two-dimensional signatures show that the minimum state aperture representation necessary to obtain robust classification with a single perspective is 20°. As previously mentioned, this is a consequence of the template heading distribution.

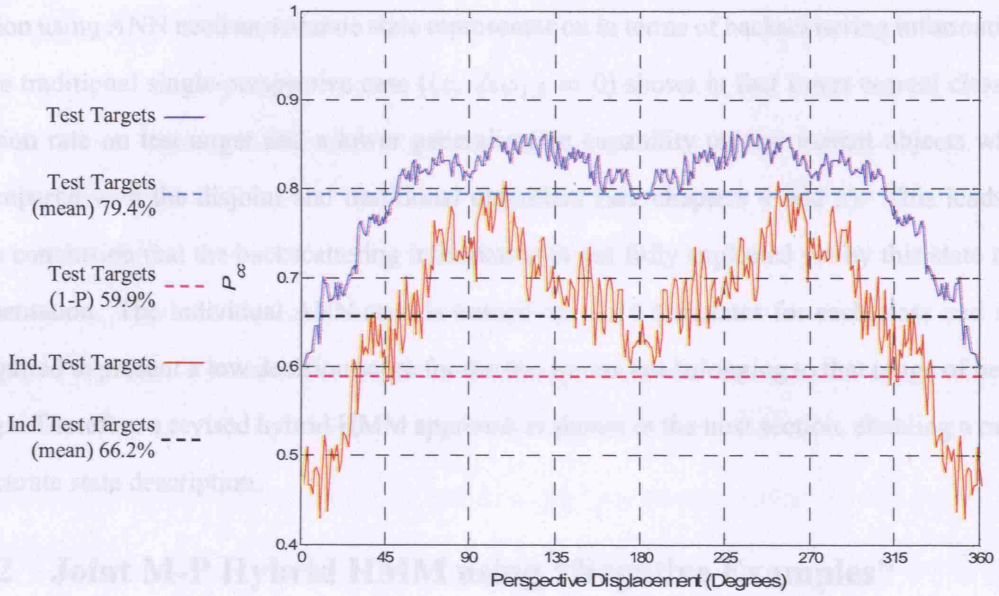

Figure 6.7: *Hybrid HMM/ANN results on MSTAR data.*

The results in a forced decision environment are shown in figure 6.7.The performance on the test targets having a representation in the training set (blue line) is compared with the correct classification rates of the target having a representation in the template set but presenting different features from the reference objects (test variant targets, red line). The single perspective (1-P) probability of correct classification for the test targets is 59.9%, a consequence of having no transition combined with the fact that the decision relies on the single perspective emissions only. The classification improvements progressively occur as the two-dimensional radar images decorrelate. When the topology $\Delta\varphi_{1,2}$ is about the same as the state aperture, a transition is measured and a rapid increase characterises the probability of correct classification. This trend is similar for the test variant targets. The main 2-P correct classification rate over all possible perspective topologies is +19.45% for the test targets and +19.55% for

the test variants, suggesting again how the second perspective significantly aids the classification and generalisation performance. As already discussed, the perspective displacement $\pi/2$ represents a good topology for the joint M-P approach.

Nevertheless, the results reported here for the joint M-P classifier are lower than the disjoint case (see Chapter 5), leading to the conclusion that HMM and their hybrid implementation using ANN need an accurate state representation in terms of backscattering information. The traditional single-perspective case (i.e. $\Delta\varphi_{1,2} = 0$) shows in fact lower correct classification rate on test target and a lower generalisation capability on test variant objects when compared with the disjoint and traditional classifiers (see chapters 4 and 5). This leads to the conclusion that the backscattering information is not fully exploited yet by this state representation. The individual ANN state is trained on $\sim$ 15 templates for each class and it is required to present a low decision score for the test inputs not belonging to that range of headings. Therefore a revised hybrid HMM approach is shown in the next section, enabling a more accurate state description.

6.2 Joint M-P Hybrid HMM using "Negative Examples"

The novel state representation is extended giving "negative" examples to expand the number of elements in the training set and improve the characterisation of each class state.

The training set for the i-th state of class given by equation 6.4 is integrated with a set of templates $\overline{\mathrm{T}}$ for which the neural network is trained to give the null score Y_{null}:

$$\overline{\mathrm{T}}(S_i) = \bigcup_{k=1}^{n_c} \overline{T}(S_i, k) \quad \text{where } \overline{T}(S_i, k) = \{\mathbf{t}_j^k\}, j \notin [\alpha_i(1), \alpha_i(2)] \tag{6.5}$$

where n_c is the number of classes, and the angular interval $[\alpha_i(1), \alpha_i(2)]$ identifies the range of headings represented by the i-th state.

The desired decision vector for the "positive" examples of equation 6.4 is an n_c-dimensional vector of zeros presenting a one in correspondence to the k-th label class represented by the particular training pattern (see Appendix C). Therefore, for the training vector set $\overline{\mathrm{T}}$, the desired output is a set of zeros $Y_{null} = [0, 0, ..., 0]$, which forces the supervised training to give a low emission. In other words, each individual state is trained to give a high

likelihood only when the input belongs to the particular set of target orientations covered.

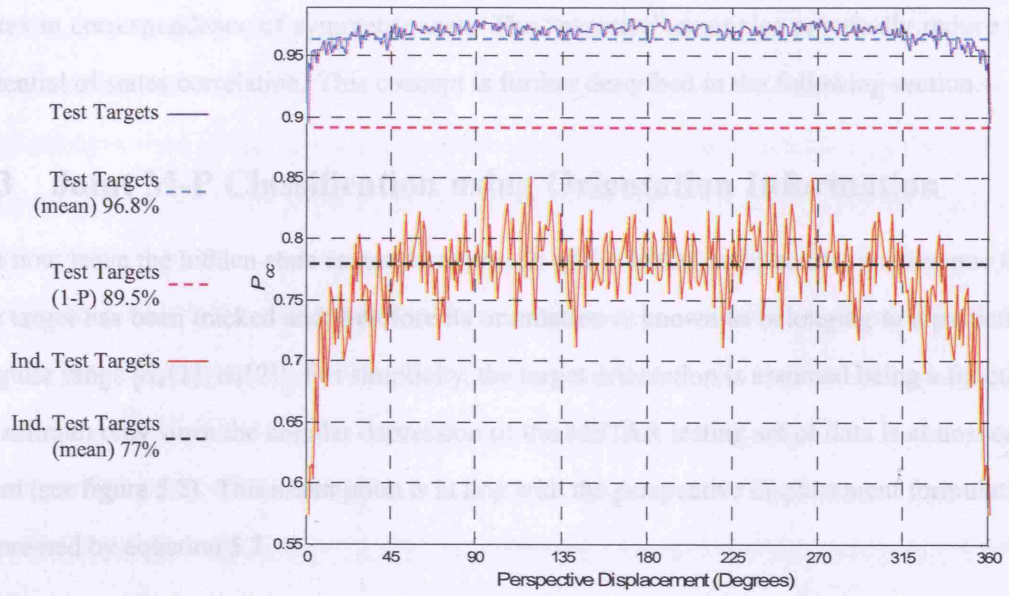

Figure 6.8: *Hybrid HMM/ANN, MSTAR Results using "negative examples" for the single state training.*

This way, the full training set is fully used for each individual state enabling the network to discriminate not only the class of the test set but also whether its orientation belongs to the range of headings covered or not. The classification performance are shown in figure 6.8 in terms of correct classification rate and generalisation capabilities.

The results are consistently improved over the hybrid HMM performance using the state representation as described in the previous section. The single perspective case probability of correct classification is 81.5%, while the corresponding generalisation capability is 57.4%. The decorrelation aspect and the correct classification rate averaged over all perspective displacements are similar to the one shown for the disjoint M-P implementation (see figure 5.20), although the overall generalisation rate (77%) is still lower than the disjoint case (81.7%). This can be explained as an effect of the negative examples. The axes of symmetry shown by ground vehicles are here not taken into account since each network is trained to recognise a single state only. Therefore, a different scattering centres distribution presented by a test

variant target at a particular heading α is more unlikely declared as generated by the corresponding class since the classifier does not consider the possible high emissions from other states in correspondence of symmetry axes. The "negative" examples practically reduce the potential of states correlation. This concept is further described in the following section.

6.3 Joint M-P Classification using Orientation Information

We now leave the hidden state sequence approach of the HMM architecture, and assume that the target has been tracked and therefore its orientation α known as belonging to a particular angular range $[\alpha_i(1), \alpha_i(2)]$. For simplicity, the target orientation is assumed being a function of azimuth only since the angular depression of the MSTAR testing set of data is almost constant (see figure 5.2). This assumption is in line with the perspective displacement formulation expressed by equation 5.2.

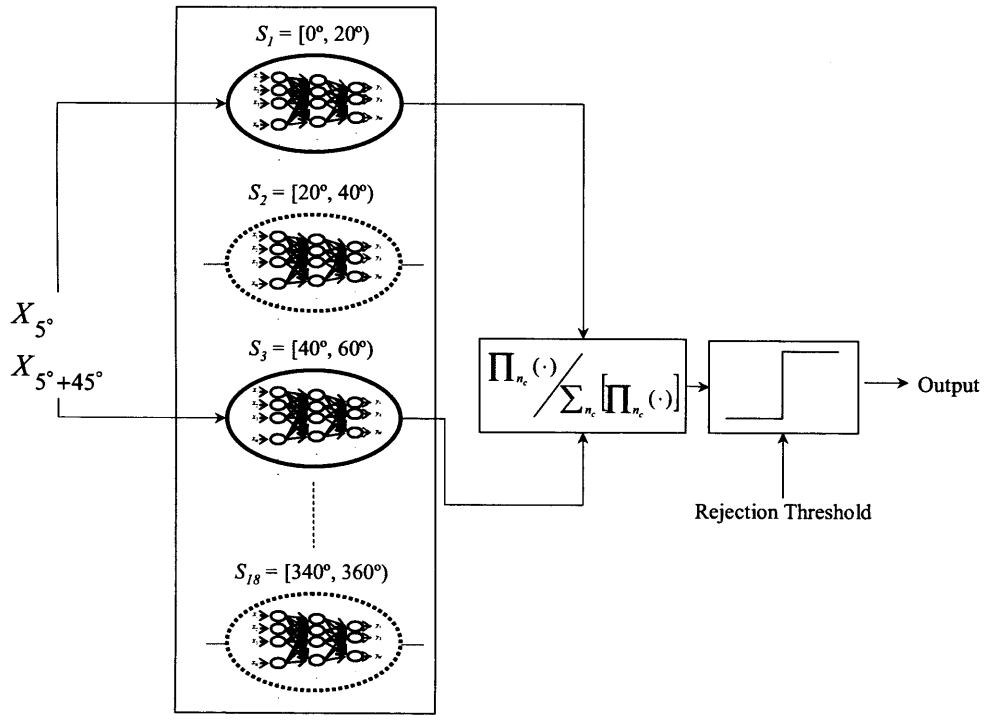

Figure 6.9: *M-P Classification using the orientation information* $\alpha' = 5°$ *and* $\alpha'' = 50°$.

Hence, the state representation can be used as described in the previous section on hy-

brid HMM using "negative" examples, where the posterior probability $P(S_i|X_\alpha)$ is calculated using neural networks.

The two collected perspectives X_1 and X_2 are associated with the corresponding orientation information α' and $\alpha'' = \alpha' + \Delta\varphi_{1,2}$. Subsequently, the two signatures are presented to the respective state networks trained on those particular headings characterising the different classes. The two activated states cover the headings displaced by the perspective disposition information $\Delta\varphi_{1,2}$. Finally, the posterior probabilities estimates are combined using the principle of Bayesian Updating as explained in Section 5.1.3.

This is illustrated in figure 6.9, where the two target orientations of the two perspectives are $\alpha' = 5°$ and $\alpha'' = 50°$. The states S_1 and S_3 are thus activated and the corresponding posterior probabilities $P(S_1|X_{5°})$ and $P(S_3|X_{50°})$ computed. Ultimately, if the combined output exceeds the rejection threshold, the declaration is made for the class presenting the highest combined posterior probability.

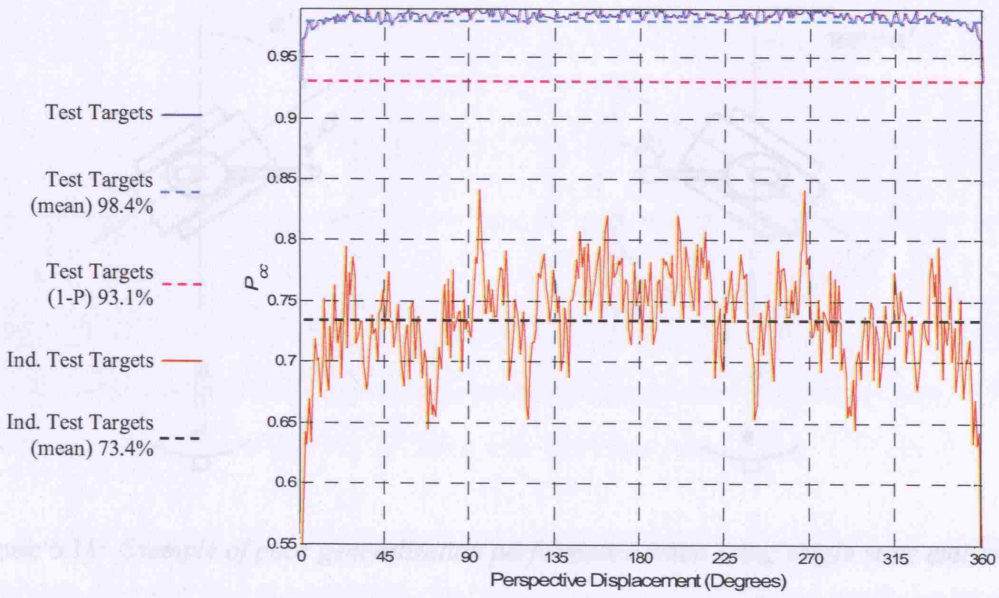

Figure 6.10: *M-P classification using the target orientation information α. Two-Perspective P_{cc} and P_{gen}.*

The results of the forced decision implementation are shown in figure 6.10. The tra-

ditional mono-perspective case shows a higher correct classification rate if compared to the traditional approach to classification where the orientation of the target is not processed. Furthermore, the Multi-Perspective probability of correct classification on test targets presents the highest average values over any two-perspective displacement (98.4%) if compared to the disjoint and joint HMM approaches.

On the other hand, the generalisation capability of the classifier is deteriorated. As already discussed in the previous section, this is a consequence of the rigid state representation using "negative" examples, which does not allow the declaration of signatures presenting different features from the ones exhibited by the training templates at a particular orientation state.

The probability of generalising test variant targets is further reduced by the activation of a single state for each perspective, and thus the discard of possible symmetries which could increase the class posterior probability in states.

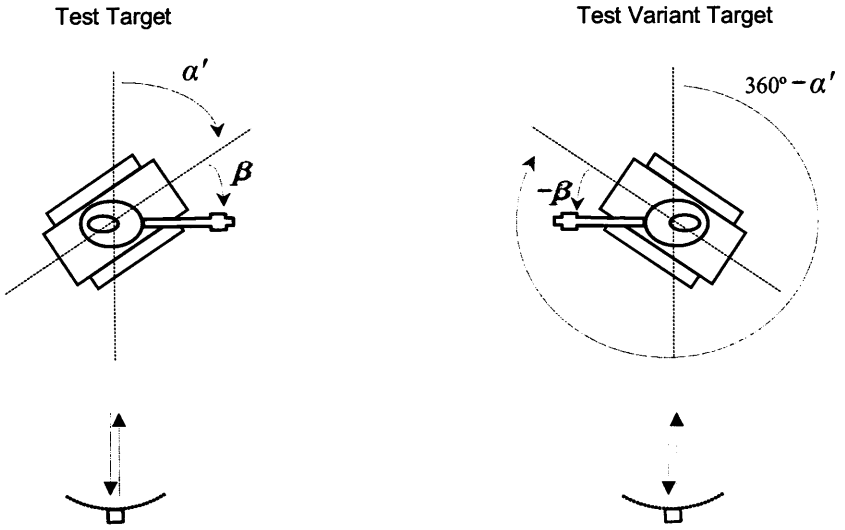

Figure 6.11: *Example of poor generalisation performance when using single state emissions for two targets, a test target (turret orientation β from the main axis) and its variant presenting (turret deviation $-\beta$).*

This can be explained by the following example. We assume the variant feature of a target being the turret position of a tank. The signature used to perform is a HRR profile. The

target having a representation in the training set (test target) for the state including the target orientation α' is on the left-hand side of figure 6.11. The turret is deviated of an angle β from the main axis orientation α'. On the left-hand side, the test variant target presents a different heading $360° - \alpha'$ and turret orientation $-\beta$ from the main axis. If the radar is assumed sensing the targets form the bottom of the figure, the range profiles result almost identical for the two different orientations, dominated by the flash of the barrel and other main scattering centres with similar spatial density distribution. Nevertheless, the states activated by the classifier are different and therefore the posterior probability of the test variant target is not supported by the state representing the heading interval $\alpha_i(1) < \alpha' < \alpha_i(2)$. This yields a reduced generalisation capability.

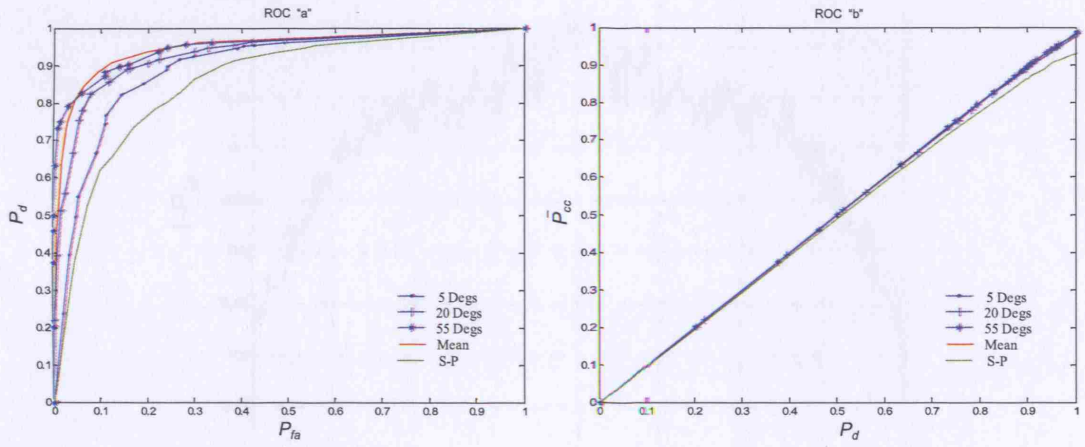

Figure 6.12: *ROC curves "a" and "b" for a single-perspective classifier (green), mean value over 360° (red) achieved with two perspectives and two-perspective performance using different angular displacements (5°, 20°, 55°) in a joint implementation.*

The ROC curves of this classification approach are now produced using different rejection thresholds. In figure 6.12, the results for the mono-perspective case (green) are shown as well as three different two-perspective displacements (blue) and the average performance over all possible perspective separations (red).

The performance improvement over the traditional single-aspect classifier can be compared with the the disjoint implementation discussed in Section 5.4 and illustrated in figure

5.18. The clear classification performance benefit of both the perspective displacement and target orientation can be evaluated through the Area Under the Curve (AUC) of both ROC curves "a" and "b" for any perspective topology.

Furthermore the reliability performance, proportional to the ratio P_{cc}/P_d of the ROC curve "b" as explained in Section 4.3.3, is significantly improved for high declaration rates also.

As analysed for the disjoint M-P implementation, the classifier requirement on the unknown targets (i.e. those targets not represented in the template library) is not to make any declaration.

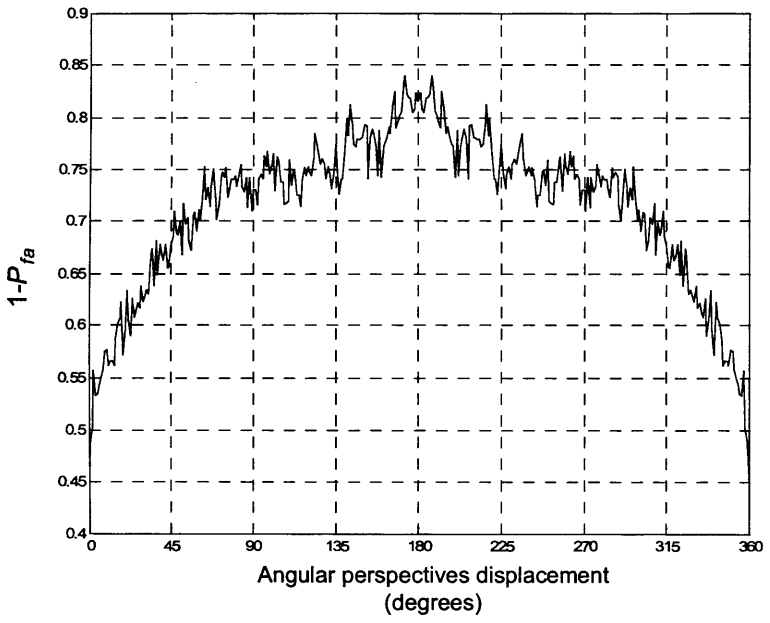

Figure 6.13: *Probability of not declaring unknown targets* $(1 - P_{fa})$ *as a function of the perspective displacement* $\Delta\varphi_{1,2}$ *for the joint classifier.*

After fixing a rejection threshold, the declaration of the unknown targets P_{fa} is measured at any perspective displacement. As illustrated in figure 6.13, the parameter $1 - P_{fa}$ increases as the perspective displacement allows for the decorrelation of the signatures. For the traditional single-perspective case, the measured false alarm rate is approximately 55%, progressively reduced to $\overline{P}_{fa} \sim 30\%$ which is the average performance over any perspective

displacement. The improvement over the disjoint classification (see figure 5.19) demonstrates that the generalisation deterioration is compensated by more significant false alarm rates. This is justified by the fact that the classifier produces a score only for the orientation given, without matching other states which could present similar features as sensed from the unknown target.

6.4 Joint M-P Classifiers using Signatures Concatenation

A number of techniques implementing the joint Multi-Perspective classifiers based on the state characterisation of the radar backscattering as previously discussed are described in this section.

The approaches are the three different methodologies already implemented for the disjoint case, i.e. when the perspective topology is not known to the system. However, the state decomposition is not used to calculate emission probabilities for the joint classifiers, but only to form a new database based for the detected perspective topology. This can be done since these classification techniques make the decision on the basis of measurements in the reference library which is interrogated every time a classification is required.

According to the detected perspective topology Φ, the template library is formed by pairs (in a 2-P problem) terns (3-P) and so forth. The size of the template does not affect the time computation of a Bayesian classifier since only the statistical parameters of the different classes are subsequently used during the decision phase. Considering a set of templates spanning 360 degrees and spaced by an angle $\mathrm{d}\varphi$, in a 2-P scenario, the *ad hoc* set of vectors forming the library is obtained by the pairs displaced by two angles:

$$\widetilde{\psi} = \left\lfloor \frac{\Delta\varphi_{1,2}}{\mathrm{d}\varphi} \right\rfloor \quad \text{and} \quad \widehat{\psi} = \left\lceil \frac{\Delta\varphi_{1,2}}{\mathrm{d}\varphi} \right\rceil \tag{6.6}$$

where $\lfloor x \rfloor$ rounds x to the nearest integer towards zero and $\lceil x \rceil$ towards infinity. If the two perspectives are displaced by less than $\mathrm{d}\varphi/2$, the library is then formed by concatenating the templates with themselves. Similarly to the HMM approach, the set of radar signatures used to form the library are organised in states. Differently, the posterior probability is not based on the emission given by a single pattern in each state but is computed considering the complete set of templates.

The posterior probability that the set of N perspectives $X_1, X_2, ...X_N$ generated by the network topology $\Phi = \{(\Delta\varphi; \Delta\delta)_{i,j}\}$ belongs to the class C_k, where $i = 1...N, j = i+1...N$, $k = i + 1...n_c$ and n_c is the number of classes in the population set is:

$$P(C_k|X_1, ..., X_N, \Phi) = \frac{P(X_1, ..., X_N|C_k, \Phi)\ P(C_k, \Phi)}{\sum_{h=1}^{n_c} P(X_1, ..., X_N|C_h, \Phi)\ P(C_h, \Phi)} \tag{6.7}$$

By maximising the likelihood $P(X_1, ..., X_N|C_k, \Phi)$ the Maximum Likelihood Estimation (MLE) of the target class C_k is computed and the class label assigned.

According to the formal description of a traditional template library given by equation 4.1, the joint M-P classifier investigated here is based on the substitution of the template library for the k-th class T_k with T'_k:

$$\mathrm{T}' = \bigcup_{k=1}^{n_c} T'_k \quad \text{where } T'_k = \left\{ \left[\mathbf{t}^k_1; \mathbf{t}^k_{1+\tilde{\psi}} \right], \left[\mathbf{t}^k_1; \mathbf{t}^k_{1+\hat{\psi}} \right], ... \right. \\ \left. ..., \left[\mathbf{t}^k_{\tau_k}; \mathbf{t}^k_{\tau_k+\tilde{\psi}} \right], \left[\mathbf{t}^k_{\tau_k}; \mathbf{t}^k_{\tau_k+\hat{\psi}} \right] \right\} \tag{6.8}$$

The element of the template set $\mathbf{t}^k_{\tau_k}$ represents the τ-th target heading signature from class K and $\mathbf{t}^k_{\tau_k+\psi}$ the corresponding signature displaced by ψ. The event C_k becomes C'_k, represented by the library T'_k.

Similarly, the test vector X_Φ is formed by concatenating the perspectives displaced by the perspective topology Φ:

$$X_\Phi = [X_1; X_2; ...; X_N] \tag{6.9}$$

By combining equations 6.9 and 6.11 with 6.7 the posterior probability of the observation sequence $X_1, X_2, ..., X_N$ is obtained as follows:

$$P(C_k|X_1, ..., X_N, \Phi) = P(C'_k|X_\Phi) = \frac{P(X_\Phi|C'_k)P(C'_k)}{\sum_{h=1}^{n_c} P(X_\Phi|C'_h)\ P(C'_h)} \tag{6.10}$$

The results of the solution to this equation are here presented for the statistical, non-parametric and multilayer neural network classifier approaches.

6.4.1 Multi-Perspective Bayesian Classifier

The Naïve Bayesian classifier is here implemented in order to evaluate the emission probabilities $P(X|S_i)$. Differently from the HMM approach we here concatenate the multiple perspectives to form the M-P pattern.

The estimation of equation 6.10 can now be treated as described for equation A.1 in Appendix A. As a consequence of the "naïve" assumption given by equation A.4 in Appendix A, there are no improvements in terms of classification: since the likelihood $P(X|C_k)$ is calculated as the product of the likelihoods $P(x_i|C_k)$ where $X = x_1, x_2..., x_m$, the result of concatenating the feature vectors is to multiply the likelihood of each element of $X_\Phi = [X_1; X_2; ...; X_N]$. This is explained in detail in Appendix E.

6.4.2 Multi-Perspective K-NN Classifier

As previously described, the set of perspectives are concatenated in order to form a single pattern X_Φ. The template library is also combined according to the detected perspective displacement Φ and the new set of reference vectors T'_k obtained for the event C_k.

The steps outlined in Appendix B are applied to the new classification problem. Differently from the Naïve Bayesian approach, the combined feature space is now effective in terms of classification.

This can be deduced from figure 6.14 where the probability of correct classification versus the number of perspectives used is plotted for the disjoint (red line) and joint (blue line) M-P implementations. The perspective topology is the same as discussed in Section 5.1.4, i.e. a perspective displacement between looks $\Delta\varphi_{i,i+1} = 20°$. Once again, as for the disjoint implementation, the most significant improvement in terms of classification can be observed when a second perspective is used. As a consequence, the backscattering description in terms of classification is enhanced by the use of the perspective topology information. The increase of correct classification rates is confirmed to be non-linear with the number of views of the target as described for the disjoint classifier. This is due to the increasing probability of presenting correlated signatures when further perspectives are processed and this does usually imply that no extra information is achieved.

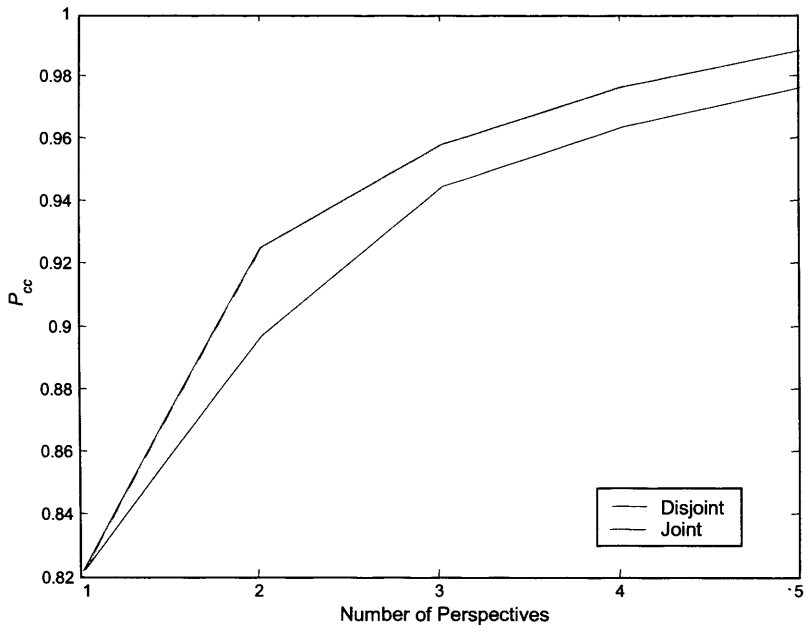

Figure 6.14: *Correct classification rates comparison between joint and disjoint K-NN classifiers where* $\Delta\varphi_{i,i+1} = 20°$.

6.4.3 Multi-Perspective Classifier using Neural Networks

The use of multilayer Neural Networks for classification has the advantage of distinguishing the learning phase – where the templates are used as examples in a supervised learning strategy in order to set the internal weight of the network – from the execution mode, i.e. the computation of the output of the NN and the subsequent decision. This aspect allows the database of templates after the learning phase to be neglected, therefore reducing the memory storage requirements and the computational burden in the execution mode. As a consequence, a bank of neural networks is trained for the available perspective topologies in the template library, formed by 72 profiles, one every 5 degrees of target rotation. This means that 72 networks are trained for the available 2-P scenarios.

As can be seen in figure 6.15, the detected perspective topology, for instance $\Delta\varphi \in (5, 10]$ degrees, activates two networks trained on $\psi_1 = 5°$ and $\psi_2 = 10°$. The partial output is subsequently combined since it represents an estimate of the posterior probabilities $P(C'_{\psi_1}|X_\Phi)$ and $P(C'_{\psi_2}|X_\Phi)$, where, differently from the representation of equation 6.11, the new set of classes

is now expressed by the template set T'_{ψ} accounting for a single topology approximation ψ:

$$\mathrm{T}'_{\psi} = \bigcup_{k=1}^{n_c} T'_{\psi,k} \quad \text{where} \quad T'_{\psi,k} = \left\{ \left[\mathbf{t}_1^k ; \mathbf{t}_{1+\psi}^k \right], \ldots, \left[\mathbf{t}_{\tau_k}^k ; \mathbf{t}_{\tau_k+\psi}^k \right] \right\} \tag{6.11}$$

The classifier is then tested with the concatenated attributes identified by the vector X_{Φ}. As for the disjoint case, the test set for a particular two-perspective scenario $\Delta\varphi_{1,2}$ is obtained by all the pairs of signatures displaced by that angle.

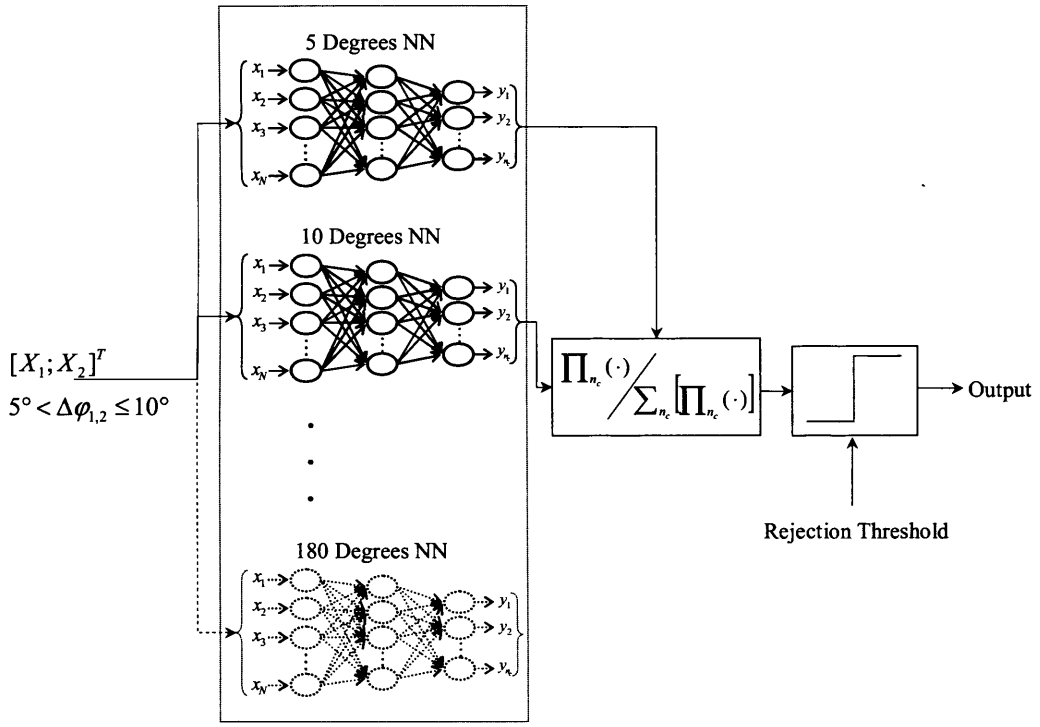

Figure 6.15: *Multi-Perspective joint classifier based on FANNs.*

The result of exploiting a progressive number of views equally spaced by 20 degrees is shown in figure 6.16. The noticeably higher performance of the FANNs classifier already discussed in Chapters 4 and 5 are her confirmed if compared with the K-NN joint classifier. Moreover, the correct classification rate improvement over the disjoint implementation is $\sim$ 0.3% when two views are used and increases with the number of perspectives. This classification improvement is limited by the particular perspective displacements considered. When other two-perspective geometries are used, the improvements given by the use of the

displacement information become more significant as can be seen from figure 6.17.

As already observed when using HMM and explained in figure 6.5, the angular displacement $\Delta\varphi = \pi/2$ is now aiding the discrimination of the classes. This is because the population of the dataset used (ADAS-MIDAS) presents ground vehicles highly separated in size, whose information is likely to appear when the perspectives sensing the target are displaced by 90 degrees. Furthermore, for $\Delta\varphi = \pi$, the classifier behaves differently depending on the class tested: for large vehicles (classes C and D), this information is highly discriminative since it automatically excludes the other two events (class A and B).

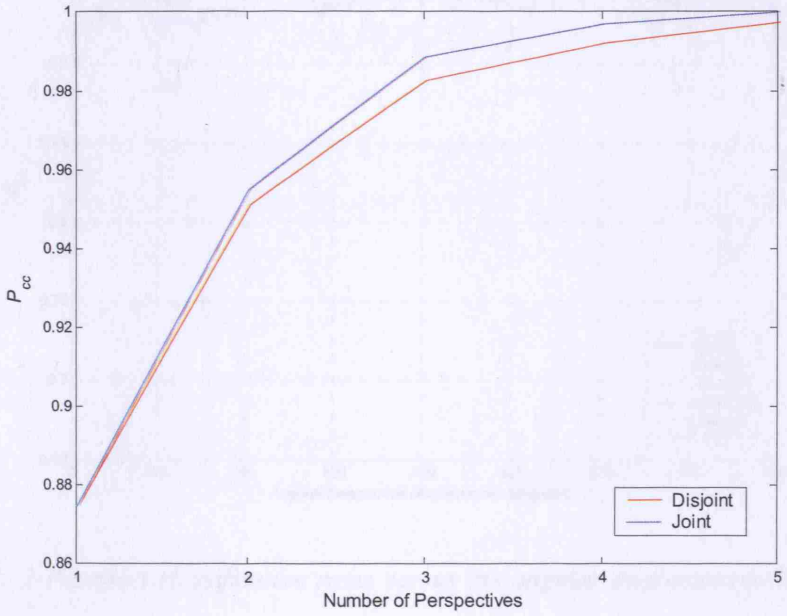

Figure 6.16: *Correct classification rates comparison between joint and disjoint FANNs classifiers where* $\Delta\varphi_{i,i+1} = 20°$.

Conversely, the smaller vehicles (class A and B) end-on profiles which can be confused with the broadside profiles of the larger classes, and the probability of correct classification is therefore reduced. Moreover an increase of information content in terms of classification can be observed for perspective displacements around 30 degrees and a significant drop is observed when 45 degrees occurs. This trend appears for the four different targets and can therefore be assumed a consequence of geometric symmetries common to the four different targets and

therefore linked to the rectangular shape of ground vehicles since it does not appear for the disjoint classification. We can therefore conclude that when the perspective displacement is used as extra information, the generalisation capabilities offered by neural networks for the target headings not represented in the training test is limited by the state representation of multiple perspectives. Besides, the optimal perspective topologies for a joint Multi-Perspective classifier are different from the disjoint implementation.

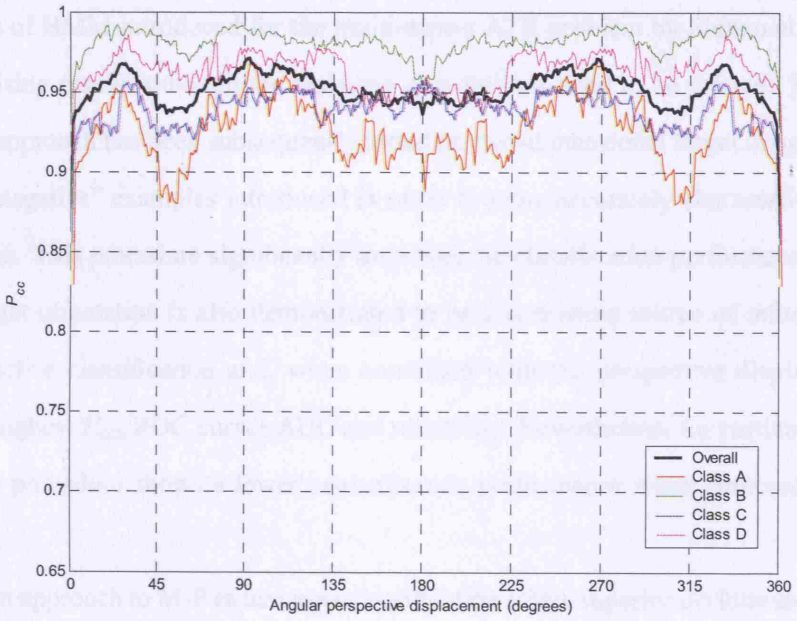

Figure 6.17: *2-P correct classification rates versus the angular displacement $\Delta\varphi_{1,2}$ for the FANNs classifier using perspective concatenation.*

6.5 Summary

In this chapter, the influence of the perspective topology on Multi-Perspective ATR is investigated using a number of classification approaches involving different techniques of perspective combination.

M-P classification has been approached using the hypothetical knowledge of the perspective displacement, which can be measured either by a network of radars or a single moving system around the target acquiring the perspective location from the estimation of the yaw,

pitch and roll components.

Similarly to the disjoint implementation, every small departure from co-location of the perspectives gives a very marked improvement in classification performance. Moreover, the increase of correct classification rate is also not linear with the number of perspectives used as already concluded for the disjoint case. Nevertheless, the perspective displacement yielding higher classification performance are different when the angular displacement is used as extra information. This is a consequence of the different target representations in the feature space.

The use of HMM introduced for the multi-aspect ATR problem by Shihao et al. (2005), shows promising results although the emission probabilities need to be refined. The Hybrid HNN/ANN approach has been subsequently tested on two-dimensional target images, and the concept of "negative" examples introduced in order to more accurately characterise the state representation. This procedure significantly improves the classification performance.

The target orientation is also demonstrated to be a consistent source of information for Multi-Perspective classification and, when combined with the perspective displacement, it exhibits the highest P_{cc}, ROC curves AUC and reliability. Nevertheless, for particular moving features, this procedure shows a lower generalisation performance when compared with the disjoint case.

The joint approach to M-P radar target classification offers superior performance over the disjoint implementation although this is counterbalanced by the increase of complexity to represent the classes in the new feature space which can be identified by concatenated templates or states in HMM. Furthermore, the difficulties of having a three-dimensional scenario (i.e. the perspectives are taken at different grazing angles) would add a further degree of complexity on the joint classification problem.

Chapter 7

Outline Structural Representation for ATR Based on Non-Radar Templates

The information content of the main scattering centres of radar signatures has been widely described and identified as the most relevant source of backscattering description (e.g. Rihaczek and Herschowitz (2000), Miller et al. (2004)). Nevertheless, as described in Section 5.2 and illustrated in figure 5.8, the correct classification rate remains stable until a particular SNR level, below which those low level scatterers disappear in noise. Amongst those scatterers, the point-like returns from the end-region at the near and far ends of the target are examined here and their relevance in ATR discussed using multiple perspectives and a structural approach.

7.1 Low Level Scatterers Information

In radar imagery, the target shape shows different features if compared with optical images of the same target. This is due to the illuminating wavelength: at X-band frequencies, fewer scatterers dominate the image obscuring the point-like ones, ultimately yielding a radar shape not easily deducible from the geometric properties of the target. For this reason, radar target recognition based on target shape features has never been encouraged (Rihaczek and Herschowitz, 2000) although it has been approached using metric algorithms (Ning et al., 2003). Shape based recognition implies a discharge of information related to the scattering centres distributed inside the contour formed by the three-dimensional surface of the target and the image projection plane.

Even if such classification might be less sensitive to a number of variable features belonging to the same class of targets that often cause misclassification, it needs to be thought as a potential support to the spatial scattering centres distribution classifiers which it could be combined with.

7.2 Radar Shape Structures

The target outline information is extracted from images obtained by coherently processing a sequence of X-band HRR profiles covering an aperture of $\omega T \simeq 4.25$ degrees. Therefore, the cross-range resolution obtained is approximately 20 cm. The resolution in range is the same as the HRR profiles. An example of (20 x 8) cm image is shown in figure 7.1 where the cross-range and range directions correspond to the vertical and horizontal axes respectively.

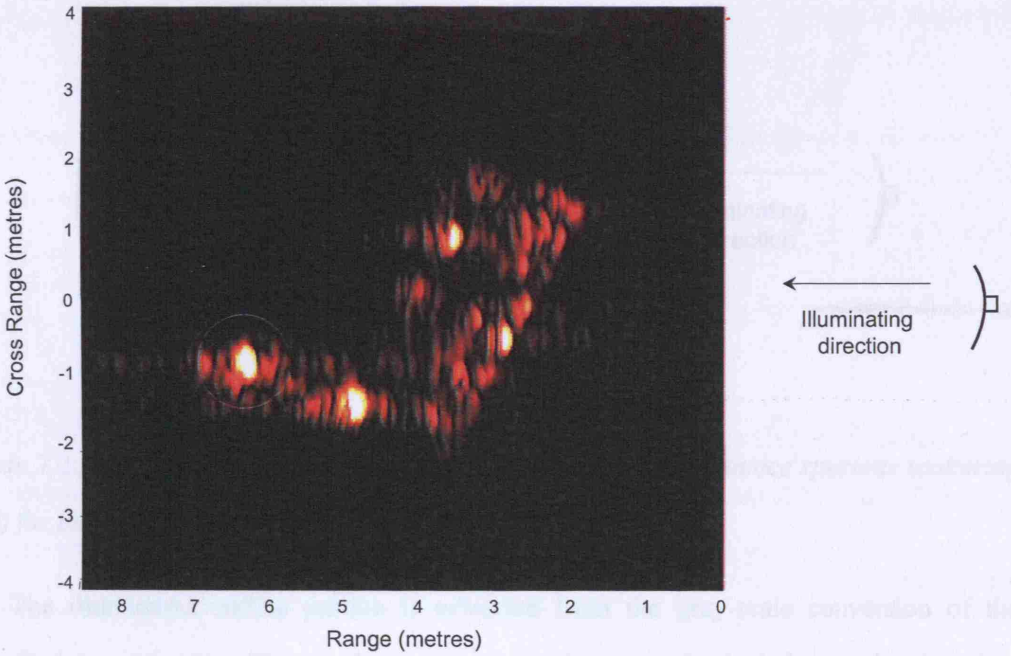

Figure 7.1: *ISAR image from ADAS data. The target exhibits a multiple bounce return highlighted (white dotted circle).*

After applying coherent two-dimensional processing, the ISAR images over small arcs of aperture are obtained. This aperture represents a particular target orientation. Since the grazing

angle is 8 degrees, the projection plane is located approximately on the plane containing the turntable. As a result, a number of scatterers belonging to the target are shadowed by the illuminated parts leading to occlusion. Therefore, the outline of the object is more detailed along the illuminated side of the target and progressively occluded on the opposite side (as shown in figure 7.1, where the radar system is located on the right side of the picture). Those scatterers presenting a higher SNR are selected to describe the illuminated shape of the vehicle for each perspective. Conversely, the strongest returns often belonging to structures located inside or outside the target perimeter. For instance, the corner-like scatterer and its multiple bounce reflection on the left side of figure 7.1, graphically explained respectively by the blue and red paths in figure 7.2, are discarded to represent the outline because not connected to the physical geometry of the target shape.

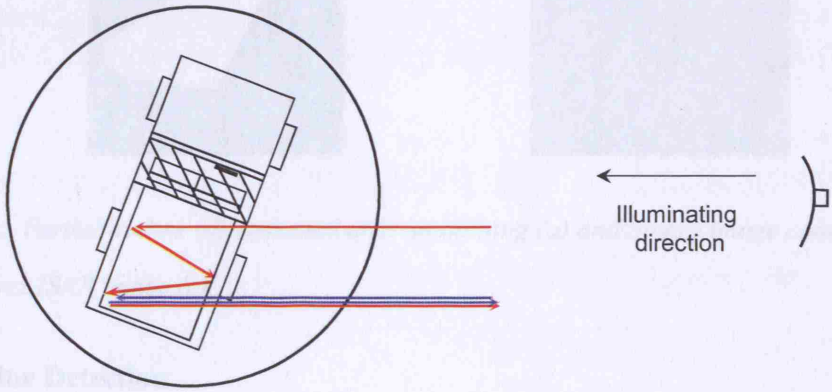

Figure 7.2: *Corner-like return (blue) and corresponding multiple bounce spurious scattering (red) for the particular target orientation of figure 7.1.*

The illuminated outline portion is extracted from the gray-scale conversion of the smoothed target image. This is obtained by assigning to each pixel the result of median filtering its neighborhood (figure 7.3-a). The corresponding binary image (figure 7.3-b) is subsequently formed after determining a luminance threshold.

The convex envelope of the target area, that is the minimal convex pixel set containing the object, is then extracted in order to highlight a shape formed by a number of straight lines (figure 7.3-c). Since the radar is sensing the target from the right side of the image,

the illuminated outline is extracted from each row of the image by detecting the first pixel transition from zero to one sliding from the right to the left of the image. The result of the algorithm applied to the ISAR image represented in figure 7.1 is shown in figure 7.3-d.

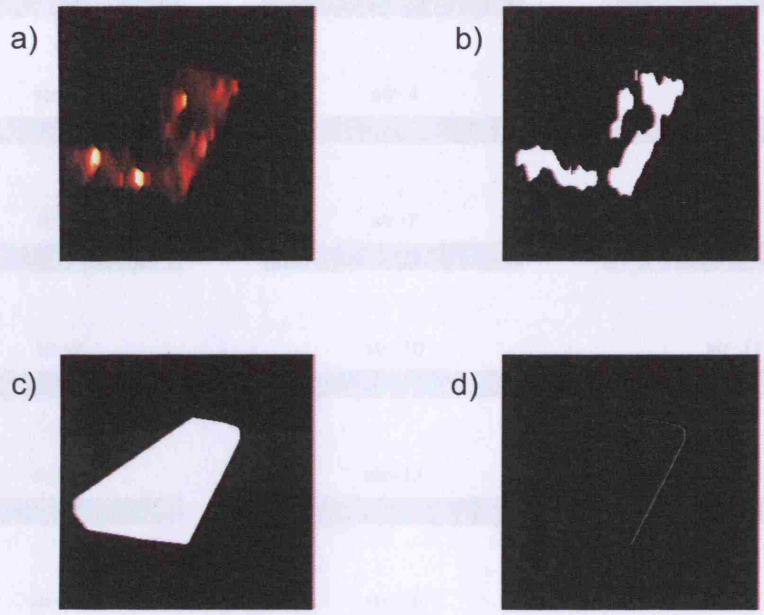

Figure 7.3: *Partial outline (d) extracted after smoothing (a) and binary image conversion (b) of the convex ISAR image (c).*

7.2.1 Line Detection

The partial outline I extracted using the procedure previously described is then processed and represented as a sequence of primitives identified by line segments ("primitives") having different slopes. The M x N image I is resolved into 8 horizontal sub-images (or regions-of-interests) $S_1, S_2, ..., S_8$ having $L = M/8$ rows and N columns. If either the first or the last row of the sub-image S_i does not present any outline pixel from the target, the sub-image is further subdivided into 3 sub-images $S_{i,1}, S_{i,2}, S_{i,3}$ then suitably padded with zero-valued pixels in order to present the same size L x N of the larger blocks. Subsequently, the line detection algorithm is applied to all elements of I denoting it as a set of *nominal* data (i.e. discrete descriptions without similarity notion (Duda et al., 2001)).

This is done by applying a kernel filtering operation (Vernon, 1991) between the partial

outline image S_i or $S_{i,j}$ and the kernel image K_j identifying the j-th structure. The kernel image presenting the highest correlation with the region of interest is finally chosen.

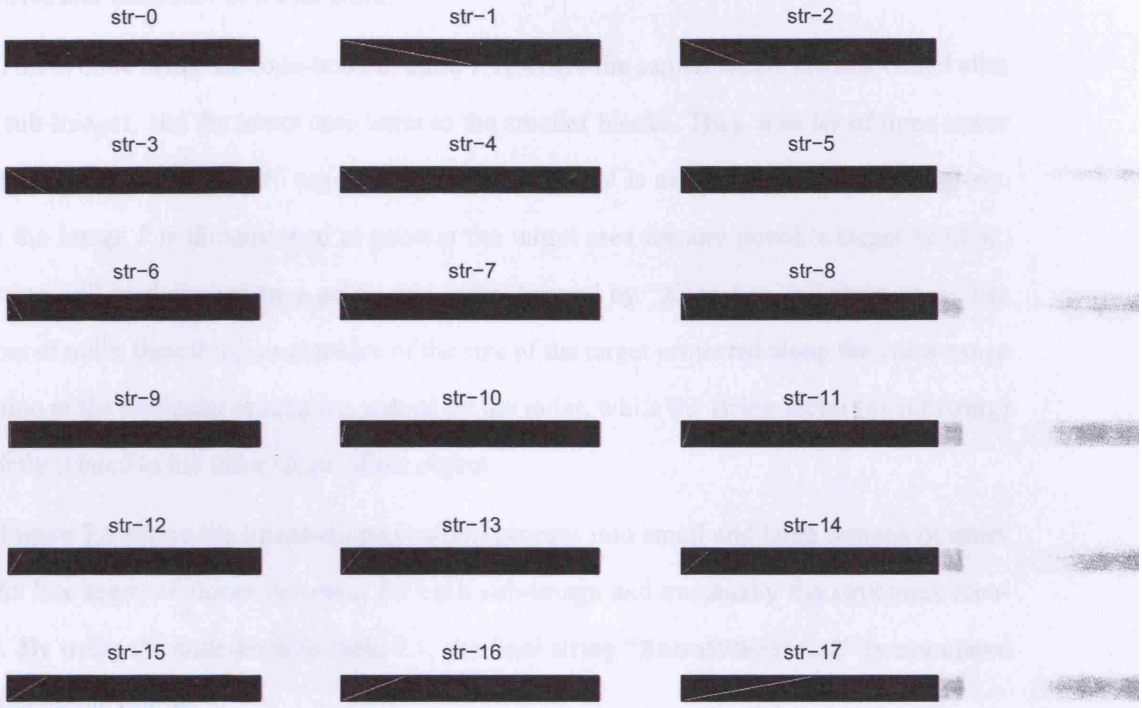

Figure 7.4: *Set of primitives forming the alphabet of structures to represent the partial target outline a sequence of straight line segments having different slopes.*

Since the line's slope domain in an image is $[0, \pi)$, assuming an angular step $\delta\alpha$, the number of primitives is equal to:

$$P = \frac{\pi}{\delta\alpha} \tag{7.1}$$

The choice of $\delta\alpha$ is dependent on the trade-off between accuracy and complexity of the outline representation since a large number of primitives yields a large codebook and, therefore, a complex model. In this work, a step $\delta\alpha = 10°$ has been used, leading to $P = 18$ structures as shown in figure 7.4.

The subdivision into more regions of the first and last sub-image encountered takes into account the more articulated partial outline sensed in those areas and, therefore, the greater number of structures presenting a smaller segment length.

7.2.2 String Representation

The image I, after being divided into a sequence of sub-images, is converted into a series of primitives and identified as a codeword.

This is done using the code-book in table 7.1, where the capital letters are associated with large sub-images, and the lower case letter to the smaller blocks. Thus, a series of three lower case letter represents a L x N region. The "Null" symbol is assigned to the empty regions. Since the image I is dimensioned to contain the target area for any possible target heading, the string will be delimited by a prefix and suffix formed by "Z" and/or "z" characters. The number of nulls, therefore, is a measure of the size of the target projected along the cross-range direction at the particular orientation sensed by the radar, while the string factor (or substring) is strictly related to the radar shape of the object.

Figure 7.5 shows the image decomposition process into small and large regions of interest, the line segment slopes detection for each sub-image and eventually the structures identified. By using the code-book in table 7.1, the final string "ZzbeMMMMmmzZ" is associated with the partial outline.

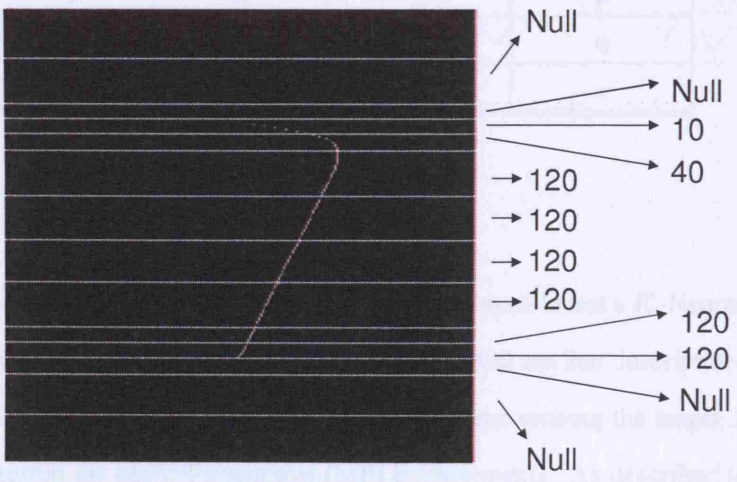

Figure 7.5: *Image decomposition into sub-images and structures identification leading to the string representation "ZzbeMMMMmmzZ".*

Table 7.1: *Regions of Interest codebook.*

		Symbol	
Structure	Line Slope	Large ROIs	Small ROIs
Null	No Line	Z	z
0	0°	A	a
1	10°	B	b
2	20°	C	c
3	30°	D	d
4	40°	E	e
5	50°	F	f
6	60°	G	g
7	70°	H	h
8	80°	I	i
9	90°	J	j
10	100°	K	k
11	110°	L	l
12	120°	M	m
13	130°	N	n
14	140°	O	o
15	150°	P	p
16	160°	Q	q
17	170°	R	r

7.3 Multi-Perspective Reconstruction

The partial outline description in terms of structures is now used to test a K-Nearest Neighbors (K-NN) classifier based on the edit distance. Since a partial outline description represented by a single radar image might be enhanced by other systems sensing the target, the classifier is subsequently tested for Multi-Perspective (MP) environments. As described in Chapter 5, the different number of nodes can be simulated by processing a series of signatures displaced by the perspective apertures described by the network topology Φ.

In information theory, the concept of similarity between nominal data can be thought of

as a metric that considers the minimum number of operations to transform a string into another.

The Hamming distance (i.e. the number of substitutions necessary to change one string into another) can only be measured between words having the same length. Since the outline from radar targets is dependent on the prominency of those major scattering centres obscuring the lower level scatterers, it is possible that less structures can be detected and the string representing the image has a variable dimension or presents a shifted sequence of structures. The Hamming distance is the upper bound of the Edit (or Levenshtein) distance (Levenshtein, 1966) which can be measured between words having different lengths and accounts for shifted versions of the same string. It is defined as the number of insertions, deletions and substitutions required to change a string into another. This is commonly implemented using a dynamic programming algorithm. For example, if we consider another hypothetical version of the target in figure 7.5 presenting a specular return from a feature far from the illuminate outline and appearing after a slight heading rotation, the resulting threshold of the gray-scale conversion will be related to a new and more intense peak and the outline size would be noticeably reduced although the structure distribution would remain similar. If we assume their structural representations as "ZzbeMMMMmmzZ" and "ZZzbeMMMmzzZ", the resulting edit distance is 3 (i.e. the deletion of the "Z", the insertion of the "M" and the substitution of the "z" into"m") while their Hamming distance is 5.

7.3.1 Target Shape Based Recognition

As explained in the Introduction, ATR can be based on reference templates either from radar measurements or Computational Electro-Magnetic (CEM) codes. The former approach is often inadequate because it is strictly related to the radar parameters and, furthermore, limited by the restricted amount of radar measurements (e.g for military applications). The latter methodology is not always reliable because of the large number of signature variability sources that make significant the difference between measured and simulated target echoes. A possible solution consists of modelling the radar backscattering from non-radar data. This approach has been investigated using the Scattering Centre Model (SCM), showing the possibility of classifying objects by means of their scattering centres spatial distributions (Miller et al., 2004).

The outline recognition can also be based on non-radar descriptions in terms of size and shape features of targets. The reference images are here deduced from the geometric target dimensions: each target heading is reproduced by rotating the physical target contour and applying the same partial outline extraction procedure previously described. The outline-based classification results shown hereafter are a gauge of the possible performance improvements if the outline structure description supports other radar backscattering approaches.

The K-NN classifier has been tested with different K values, showing the highest Correct Classification Rate (CCR) when $K = 3$ was used. The targets forming the population problem are the three ground vehicles belonging to the ADAS dataset. The set has not been expanded using the MIDAS data since the correct physical size of the vehicles was not known.

The relative confusion matrix (table 7.2) reveals the possibility of discriminating shape structures from partial information. Nevertheless, a $P_{cc} = 66.7\%$ is poor and can be explained as a consequence of the reduced amount of information used for describing the target contour. As observed in the previous chapters, the correct classification rates for a two-perspective classifier depend on the angular displacement $\Delta\varphi_{1,2}$ between the views sensing the target, although the benefit from employing multiple signatures are achieved for any topology Φ allowing for enough decorrelation between the signatures.

Table 7.2: *Partial outline classification results using a single view of the target ($P_{cc} = 64.7\%$, Average reliability $R_{avg} = 66.7\%$).*

		Output $^{(Reliability)}$ (%)					
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)
A	(360)	**63.9**	(53.6)	4.8	(5.8)	31.3	(32.1)
B	(360)	16.9	(14.1)	**73.5**	(88.4)	9.6	(9.9)
C	(360)	38.6	(32.3)	4.8	(5.8)	**56.6**	(58)

Nevertheless, as explained in Chapter 6, there are favourable perspective displacements depending on the population set and the Multi-Perspective approach used (i.e. using the per-

spective displacement information or otherwise).

The results of the two-perspective classifier using $\Phi = \{\Delta\varphi_{1,2} = 80°\}$ in table 7.3 show a probability of correct classification and reliability increase if compared with the single-view performance, respectively +8.8% and +8.6%.

Table 7.3: *Partial outline classification results using two perspectives,* $\Phi = \{\Delta\varphi_{1,2} = 80°\}$ *(*$P_{cc} = 73.5\%$*, Average reliability* $R_{avg} = 75.3\%$*).*

		Output $^{(Reliability)}$ (%)					
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)
A	(360)	**84.3**	(60.9)	2.4	(2.6)	13.3	(19.4)
B	(360)	4.8	(3.4)	**88**	(94.8)	7.2	(10.4)
C	(360)	49.4	(35.7)	2.4	(2.6)	**48.2**	(70.2)

Although the increase of classification performance over the single perspective case, the correct classification rates of the target class C are deteriorated. This could be misinterpreted if not linked to the increase of reliability. Since it is a forced decision implementation, the classifier is always declaring. This means that when two perspectives are used, the classifier decides for class C more reliably, and it is therefore more likely that when a target of class C is declared it is actually that target the input to the classifier.

Table 7.4: *Partial outline classification results using two perspectives and the topology information (joint mode),* $\Phi = \{\Delta\varphi_{1,2} = 80°\}$ *(*$P_{cc} = 77.1\%$*, Average reliability* $R_{avg} = 79.1\%$*).*

		Output $^{(Reliability)}$ (%)					
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)
A	(360)	**68.7**	(72.2)	–	(–)	31.3	(26)
B	(360)	4.8	(5)	**84.3**	(100)	10.9	(9)
C	(360)	21.7	(22.8)	–	(–)	**78.3**	(65)

The results of the K-NN classifier implemented using the perspective displacement as described in Section 6.4.2 are shown in table 7.4. Both reliability (+3.8%) and correct classification rates (+3.6%) are improved through the use of the perspective topology.

Table 7.5: *Partial outline classification results using three perspectives and the topology information (joint mode),* $\Phi = \{\Delta\varphi_{1,2} = 30°, \Delta\phi_{2,3} = 30°\}$ *(*$P_{cc} = 78.9\%$*, Average reliability* $R_{avg} = 81.4\%$*).*

		Output $^{(Reliability)}$ (%)					
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)
A	(360)	**65.1**	(78.3)	–	(–)	34.9	(26.2)
B	(360)	6	(7.2)	**83.6**	(100)	10.4	(7.8)
C	(360)	12.1	(14.5)	–	(–)	**87.9**	(66)

By adding a further perspective in the network, the three-perspective classifier is simulated. The perspective topology showing a significant correct classification rate improvement is achieved when $\Phi = \{\Delta\varphi_{1,2} = 30°, \Delta\phi_{2,3} = 30°\}$. These perspective displacements are used to test the 3-P classifier. The classification performance of a three-perspective classifier using the angular displacement information is reported in table 7.5. The correct classification rates achieved shows a negligible increment of recognition capabilities, suggesting that the upper bound of the outline structural representation by means of multiple perspectives is reached.

7.4 Summary

A novel radar target description has been introduced in terms of outline structures decomposition from 2-D high resolution images: the illuminated side of the target is detected and transformed into a set of straight lines from the alphabet of primitives. The non parametric K-NN classifier has been implemented using a template library from non-radar features, showing the separation capabilities on real turntable data.

Although this methodology is more suitable for supporting other approaches using the main scattering centres information, this recognition procedure might show valuable classification performance when integrated in a Multi-Perspective environment. Furthermore, because of the need for connecting partial outline structures, the knowledge of the perspective topology contributes for further outline separation and, therefore, for a more accurate outline structural representation.

Chapter 8

Multi-Frequency and Multi-Perspective Classification

The information content of radar target signatures is the key aspect for Automatic Target Recognition (ATR). In this Chapter, the role of high range resolution is investigated in relation with the influence of the illuminating wavelength on the backscattering description. The classification performance is evaluated using (i) full scale 2-D Inverse Synthetic Aperture Radar (ISAR) images obtained from Stepped-Frequency compression and (ii) the corresponding sub-spectra of the target reflectivity function forming lower resolution images at different centre frequencies. Therefore, the classification performance comparison between the two different techniques is made on the same data processed in different ways. The classification rates given by different combinations of RFs are also evaluated and compared with the coherent reconstruction of the full bandwidth. Furthermore, the classification results are analysed using multiple aspects sensing the target.

8.1 Frequency Dependency of Target Backscattering Information

High Range Resolution (HRR) is achieved by synthesising large bandwidths. The information content of the received signal modulated by the target resides not only in the spatial discrimination of the scatterers identifying the target, but also in the scattering behaviour of the radar object when illuminated by a set of different frequencies. In this chapter, the information content of radar target signatures is analysed in terms of range resolution (coherent bandwidth

reconstruction), Multi-Frequency (M-F) illumination and the corresponding classification performance evaluated using multiple aspects collected by a network of radars or by an individual system sensing the target from different perspectives.

In optical images, the illuminating waveform has a wavelength much shorter than the dimension of the elementary structures forming the object. Therefore, optical targets appear as a set of point-like scatterers with the only exception of rare specular reflections. Furthermore, the reduced incidence of end-region scattering behaviours means that the scattering centre radiation pattern will not show pronounced sidelobes. As a result, the optical image is formed by a set of point scatterers which resemble the physical spatial distribution of the surface roughness and discontinuities of the target. The progressive reduction of the illuminating frequency to the radar range (1-40 GHz) implies different scattering behaviours of manmade targets. In the radar optics region, for instance, when the illuminating wavelength λ becomes much smaller than the target size D, radar objects can be seen as a set of independent elementary scatterers coherently contributing to the target radar signature.

The simplest canonical shape is the point-like scatterer, characterised by an omnidirectional and uniform radiation when not occluded by other parts of the target. Although ideal, at very high frequencies, the discontinuities on the target surface cause abrupt variation of the induced currents and, therefore, the appearance of the low level scatterers in the radar signature. These scatterers often satisfy the point-like properties. Other scattering mechanisms - such as specular reflection, multiple bounces, edge diffraction, end region effect, and surface waves - characterise a number of canonical shapes. These primitives present peculiar reradiation patterns and an RCS dependent on their physical dimension and structure (Knott et al., 1985).

For instance, a corner reflector is dominated by multiple reflections and its RCS is in direct proportion with the square of the illuminating frequency. The maximum RCS from a cylinder occur when the incident field is normal to its axis and present a linear proportion with the illuminating frequency. Some scattering shapes show an inverse proportion with frequency. Therefore, if a range of frequencies is spanned to sense the elementary shape, the information gained by the corresponding RCS could be used to discriminate the dominant

scattering mechanism. Furthermore, a number of shapes will appear only in a particular set of illuminating frequencies. This concept could be extended to complex targets.

The RCS of complex targets is also highly sensitive to both the incident and scattering directions. This is due to a number of factors responsible for the radar signature variability, and as a consequence of the ATR performance reliability deterioration as explained in Chapter 4, which could be overcome by using multiple uncorrelated signatures from the target, i.e. using different perspectives. The combination of frequency and perspective diversity is also discussed here.

8.2 Pattern Description and Classification Approach

ISAR images are formed from the previously described ADAS and MIDAS data, using a sequence of HRR profiles from turntable measurements. The profiles are generated from an X-band ($\sim$ 10 GHz) radar having an instantaneous bandwidth of 500 MHz, transmitting Stepped-Frequency LFM chirps. The turntable dataset is first prepared by removing any stationary clutter to avoid the classifier to exploit this consistent information (Showman et al., 1998).

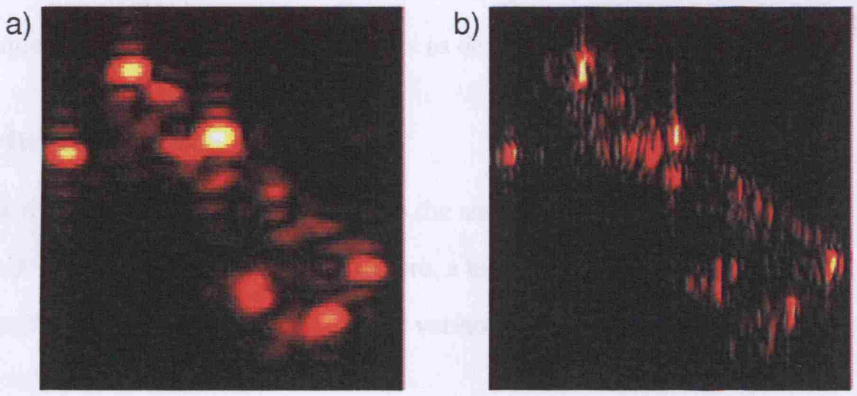

Figure 8.1: *Fixed cross-range resolution ($\Delta r_c = 30$ cm) ISAR Images from a lorry. (a) single chirp $\Delta r_s = 30$ cm, (b) Stepped-Frequency Chirp $\Delta r_s \sim 8$ cm.*

The pattern to be classified is deduced from 2-D images obtained by covering apertures $\psi = 2.83$ degrees. Therefore, the cross-range resolution obtained is approximately 30 cm for every dataset used. An example of (30 $\times$ 30) cm image and the corresponding higher slant

range resolution are shown in Figure 8.1.a where the cross-range and range directions can be identified with the vertical and horizontal axes respectively and the radar system is located at the right-hand side of the figure. The full set of images at eight different centre frequencies are later illustrated and investigated in figure 8.3.

In Figure 8.1.b, ISAR images covering the same aperture ψ are shown at higher range resolution, synthesizing a large bandwidth by using a number of chirps stepped in frequency. It is clear the higher separation of multiple scatterers along the down-range direction, as well as the aid that could be provided to automatic recognition. The six targets, previously described in Chapter 4 are organised into four classes (A, B, C, and D), plus a confusing class (i.e. not used to train the classifier). The last data file (D_{ind}) corresponds to a test vàriant target, a different version of the target used to represent class D in the training set.

The training set is made up of a sequence of images for each class at specific target headings. As previously described in Section 4.4.4, 32 images spaced by 11 degrees form the template library for each of the four classes, while the test set consists of 350 images spaced by 1° from the test target set and 217 from the test variant and unknown classes. The classification approach used involves a first PCA stage, followed by FANNs classifiers implemented using both a single perspective and multiple aspects as described in Chapters 2, 4 and 5.

8.3 Multi-Frequency Classifier

The target reflectivity reconstruction requires the assumption that the target is not moving during the time length of the chirp burst. Therefore, a high PRF is needed between consecutive RF signals. Furthermore, the scattered signature variability due to the speckle may be significant and the reflectivity reconstruction distorted.

As mentioned in Chapter 1, the advantages in of wideband transmitted signals are the high resolution achievable and the information content of the backscattering at different frequencies while the instantaneous bandwidth requirement is kept constant. The first two aspects will be examined in terms of classification performance.

The multiple RF received signatures, as shown in Figure 8.2, can be considered as separate patterns to be given as input to multiple classifiers tuned to the different centre frequencies.

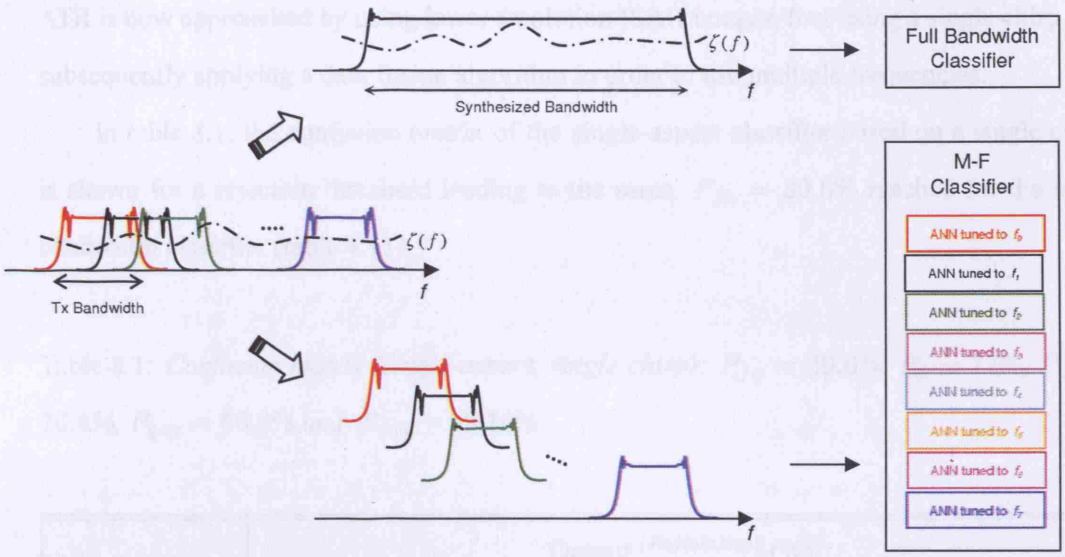

Figure 8.2: *Reflectivity reconstruction and Multi-Frequency classification approaches.*

Although the non-exploitation of the higher range resolution and the consequent classification performance deterioration, the M-F approach will allow a more robust and fast processing. Moreover, the resolution obtained by the synthesized wideband is reduced by the clipped side portions of the sub-spectra since the chirps are overlapping (see equation 1.15), whereas if single chirps are processed, the sub-spectra are fully utilised.

8.4 Single-aspect - large bandwidth reconstruction results

The classification performance of the classifier using the full synthesized bandwidth obtained by Stepped-Frequency processing the RF chirps is shown in Table 4.11 (Chapter 4). The confusion matrix is obtained after fixing a rejection threshold that yields a $P_{fa} = 39.6\%$.

The target D_{ind} gives a measure of the generalisation capabilities of the classifier. The probability of declaration is measured as the mean value of $(1 - \sum P_{miss})$ calculated for the classes used to train the classifier and the test variant targets. P_{fa} is measured on the "unknown" target only. It is worth noting that the "unknown" target presents similar shape features to the target of class A, with which it is more likely to be misidentified.

8.5 Single-aspect - Multi-Frequency lower resolution results

ATR is now approached by using lower resolution ISAR images, first using a single chirp and subsequently applying a data fusion algorithm in order to use multiple frequencies.

In table 8.1, the confusion matrix of the single-aspect classifier based on a single chirp is shown for a rejection threshold leading to the same $P_{fa} = 39.6\%$ reached for the large bandwidth classifier (table 4.11).

Table 8.1: *Confusion matrix (single-aspect, single chirp):* $P_{fa} = 39.6\%$, $P_d = 77\%$, $P_{cc} = 76.4\%$, $P_{gen} = 56.2\%$ *and* $R_{avg} = 99.16\%$.

		Output $^{(Reliability)}$ (%)								
Input	(n_s)	A	(R_A)	B	(R_B)	C	(R_C)	D	(R_D)	*Unknown*
A	(350)	**67.8**	(98.8)	1.1	(1.4)	–	(–)	0.3	(0.4)	30.8
B	(350)	0.8	(1.2)	**78.1**	(98.6)	–	(–)	–	(–)	21.1
C	(350)	–	(–)	–	(–)	**84.7**	(100)	0.3	(0.4)	15
D	(350)	–	(–)	–	(–)	–	(–)	**75.1**	(99.2)	24.9
D_{ind}	(217)	–		17.1		0.5		**56.2**		26.2
Unknown	(217)	35.9		3.2		0.5		–		60.4

As expected, the use of reduced range resolution deteriorates the declaration probability P_d (from 95.7% to 77%) as well as the correct classification rate P_{cc} (from 93.9.7% to 76.4%) and the generalisation capability of the network P_g (from 62.7% to 56.2%). Nevertheless, the figures reported for the lower resolution case qualitatively reflect the high resolution results: the "unknown" target class is mostly confused with class A and the test variant target D_{ind} frequently not declared or misidentified with class B. The correct classification rates for class A are the lowest between the test targets, suggesting similar decision boundary characteristics in the feature space. The reliability at the rejection threshold fixed are also comparable using the two different approaches, suggesting that the classifiers are unbiased. In particular, the

decisions made for target class B and class D show a slightly reduced classification reliability for both the coherent reflectivity reconstruction and the M-F approaches.

The output of the non linear mapping of the ANN classifier is an N-unipolar cube $[0,1]^N$ where N is the number of classes forming the population. The decision is made by sensing the greatest component of the output vector which can be therefore considered as a measure of the confidence of the decision. The input pattern of the j-th network is a vector representing the target backscattered signature X_{f_j} where f_j is the illuminating centre frequency in the available set of illuminating frequencies $F = \{f_1, ...f_n\}$. Different combinations of centre frequencies $F' = \{f_i : i \in \Omega\}$, where Ω is a subset of the available chirps, can be adaptively chosen in order to maximise the classification performance. The fusion algorithm consists of updating the posterior probabilities of the networks tuned on different centre frequencies. In Figure 8.3, the (30 $\times$ 30) cm resolution ISAR images from a Tank at eight different centre frequencies at X-band are shown. Cross range is along the vertical axis, while slant range can be approximated with the horizontal one. As a consequence of the small grazing angle (8°), a portion of the target on the right-hand side is locally shadowed.

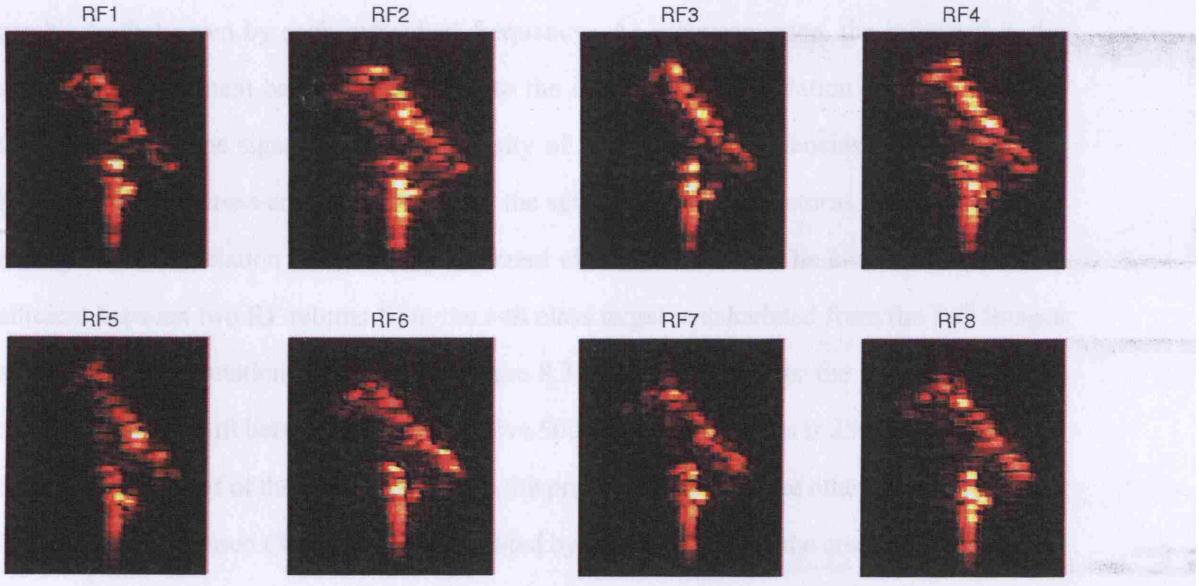

Figure 8.3: *Lower resolution ISAR images from a tank at different centre frequencies.*

At that particular heading, the target shows a specular reflection from the barrel. This scattering behaviour is independent of the incident frequency and appears for the entire set RF chirps. The point-like scattering from the illuminated part of the target is also frequency invariant and the partial outline can be delineated. A number of scattering centres located inside the tracks and geometrically characterised as a set of corner-like shapes, appear in different positions when the centre frequency changes. The same effect can be seen at the junction between the turret and the barrel, where a scattering centre appears with different intensities depending on the illuminating wavelength. This phenomenon can be traced back to a small speckle component as well as the backscattering frequency dependence on the illuminating frequency when complex targets are sensed. As a consequence, certain centre frequencies highlight different features, giving rise to frequency dependent classification results and reliability. Particularly, the scattering centre appearing at the beginning of the track opposite the barrel is clear for RF3 and RF4 (i.e. chirps 3 and 4) only. This may suggest a resonance behaviour at those wavelengths which can be associated with those particular frequencies. If important for classification the choice of frequency will clearly have a significant effect on the resulting performance.

If two signatures are highly correlated, the total information brought by the pair is comparable to that given by each individual frequency. As a consequence, the information description improvement has proportionality to the degree of de-correlation achieved. Hence, the difference in the signature due to diversity of illuminating frequencies can be estimated by measuring the cross-correlation between the set of available RF returns. In figure 8.4, the average cross-correlation over the four different classes is shown. The cross-correlation coefficient between two RF returns from the i-th class target is calculated from the 2-D images at same target orientation (as shown by figure 8.3) and averaged over the 360 degrees. The centre frequency shift between two consecutive 500 MHz chirp returns is 250 MHz. Therefore they overlap for half of their bandwidth with the previous pulse and the other half with the one following (as discussed Chapter 1 and illustrated by figure 1.6). Thus the cross correlation progressively decreases as the two RF shift increases. The mean cross-correlation between RF1 and RF8 is 0.78. This value shows that the information content of the two centre frequencies

returns spaced by 1.75 GHz, although still correlated, can be used to provide different target information to improve classification.

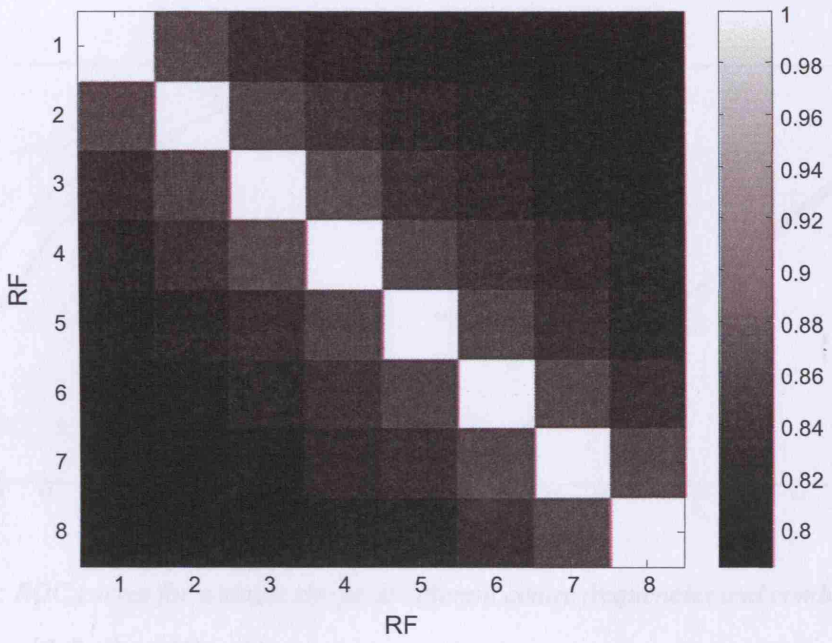

Figure 8.4: *Cross-correlation between consecutive RF returns from 500 MHz spaced by 250 MHz. The results are averaged over the four classes, and over the 360 degrees of aspect angles.*

This conclusion is verified by a mean cross-correlation value of 0.62 between random signatures in the database, representing the lower bound of information similarity between signatures. It is therefore concluded based on these measurements and observations that there will be an improvement in classification performance when multiple chirps are used. Particularly, when the returns are not from adjacent chirps, the information content of the M-F pattern is expected to be maximised.

In figure 8.5, the ROC curves of single RFs are compared with different combinations of chirps. As anticipated, the classification performance given by particular frequencies appear more robust than others (larger area under the curves). This is due to the RCS of those features that become more prominent at different illuminating frequencies, making the targets more

easily separable in the feature space. Furthermore, a significant classification improvement is achieved by exploiting different RFs, for instance using a combination of three frequencies (magenta line).

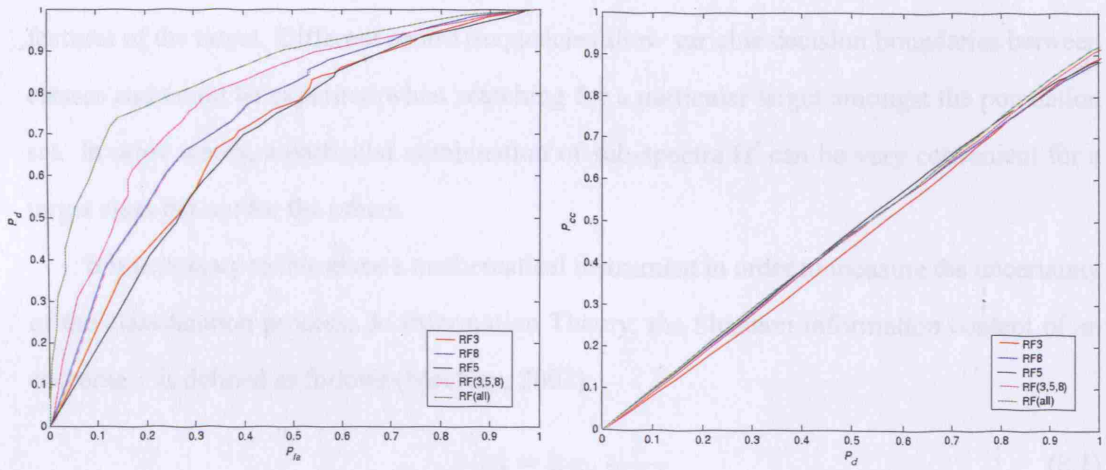

Figure 8.5: *ROC curves for a single chirps at different centre frequencies and combinations of chirps:* $\Omega = \{3, 5, 8\}$ *and* $\Omega = \{1, ..., n\}$.

Each chirp is stepped in frequency by 250 MHz, therefore for the particular combination used, where $\Omega = (3, 5, 8)$, the frequency separations $\Delta f_{3-5} = 500$ MHz and $\Delta f_{5-8} = 750$ MHz occur. This combination shows the highest area under the ROC curves when three chirps are used, due to the non overlapping spectra. Conversely, when adjacent chirps are used, the information content processed leads to less significant classification improvements. By using the complete set of sub-spectra, the M-F classifier exhibits the highest classification performance (green line). This is a consequence of processing the entire target reflectivity in the frequency domain.

The reliability R_{avg}, as discussed in Chapter 4, is proportional to the ratio between correct classification and declaration rates, thus and it can be deduced from the ROC curve 'b'. The reliability progressively increases with the number of frequencies used, meaning that the decision score is consistently improved by the Bayesian updating processing.

As summarised in the next Subsection, bandwidth to be equal, the use of the coherent

spectrum reconstruction approach yields higher classification performance if compared with the M-F methodology.

8.5.1 Adaptive Frequency Selection

The reflectivity function features drastically vary depending on the geometric and electric features of the target. Different centre frequencies show variable decision boundaries between classes and could be exploited when searching for a particular target amongst the population set. In other words, a particular combination of sub-spectra Ω' can be very convenient for a target class but not for the others.

It is necessary to introduce a mathematical instrument in order to measure the uncertainty of the classification process. In Information Theory, the Shannon information content of an outcome x is defined as follows (MacKay, 2003):

$$h(x) = \log_2 \frac{1}{P(x)} \tag{8.1}$$

where $P(x)$ is the probability of the event x. Considering the discrete random event X being the outcome of a classifier with possible outcomes x_i taking on values $1, ..., n_c$ where n_c is the number of classes in the population set, its entropy is defined as the average Shannon information content:

$$H(X) = \sum_{i=1}^{n_c} P(x_i) \log_2 \frac{1}{P(x_i)} = -\sum_{i=1}^{n_c} P(x_i) \log_2 P(x_i) \tag{8.2}$$

Given the confusion matrix for a subset $S = \{S_A, S_B, S_C, S_D\}$ in the training set T, the entropy $H(S_i)$ can be thought of as a measure of the uncertainty of the classifier for the i-th class. After processing the confusion matrices corresponding to different centre frequencies it is possible to determine for each class i a scale the optimum RFs to minimise the relative entropy $H(S_i)$.

In figure 8.6, the resulting entropy using eight Stepped-Frequency chirps is shown. The entropy is calculated as the average $H(S_i)$ measured at different rejection thresholds, in order to obtain the uncertainty level unbiased by the particular operational condition.

It can be observed that when the application priority is to recognise a target class A, the minimum entropy for that class has to be sought. This is verified by sensing the target

and collecting its reflectivity at RF $= 8$. Conversely, when the discrimination of target B is required, the entropy is minimised at a lower freqeuency (RF $= 3$).

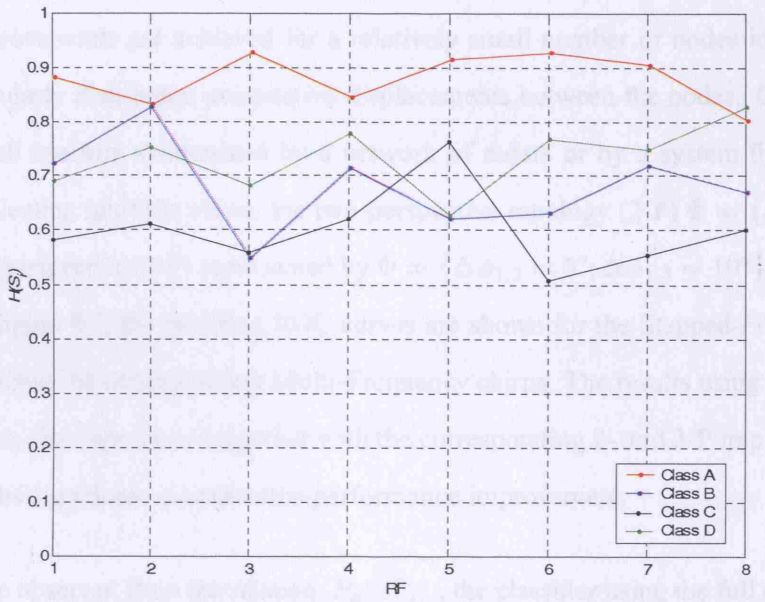

Figure 8.6: *Single class entropy $H(S_i)$ computed from a set of training samples S obtained by illuminating the targets with different centre frequencies.*

This can be a means of further refining the classification performance also when the M-F is applied, by selecting the combination of frequencies that guarantee the minimum entropy for a particular target class. For instance, to enhance the recognition capabilities for the target class D, the optimal choice for a 3-F classifier would be a selection of centre frequencies $\Omega = 1, 3, 5$.

The class entropy described so far has been deduced from the test targets only, although it could be extended to test variant targets and applied to the "unknown" class. The oscillating trend of the single class entropy suggests the transient behaviour of certain scattering mechanisms in the frequency domain.

8.6 Multi-Frequency versus Large Bandwidth and Multiple Perspectives

The analysis of M-P classification discussed in Chapter 5 shows that the most significant classification improvements are achieved for a relatively small number of nodes in the network and for particularly distributed perspective displacements between the nodes. Considering a relatively small aperture synthesized by a network of radars or by a system flying past the target and collecting multiple views, the two-perspective topology (2-P) $\Phi = \{\Delta\phi_{1,2} = 5°\}$ and the three-perspective (3-P) represented by $\Phi = \{\Delta\phi_{1,2} = 5°; \Delta\phi_{1,3} = 10°\}$ are hereafter analysed. In figure 8.7, the resulting ROC curves are shown for the Stepped-Frequency processed images and the corresponding Multi-Frequency chirps. The results using a single view of the target are also reproduced together with the corresponding 2- and 3-P implementations, highlighting the significant classification performance improvement.

As can be observed from the relation $P_d - P_{cc}$, the classifier using the full reconstructed bandwidth exhibits a higher rate of recognition on what has been declared. On the other hand, the M-P environment shows its effectiveness on the relation between P_{fa} and P_d : the classification performance improvement is significant for both the reconstructed bandwidth and the M-F approaches when an unknown target is sensed. Moreover, for applications where low P_{fa} are tolerated, the M-F approach guarantees comparable declaration rates and average reliability with the coherent reconstructed reflectivity case. The classification results corroborate the importance of range resolution in terms of classification and the Multi-Perspective effectiveness independence of the resolution used. The adaptive selection of the subset Ω of centre frequencies on the basis of the particular object sought in order to increase the likelihood of its recognition can be applied to the selection of the topology Φ. This can be done when the multiple perspectives can be chosen such as in airborne radars flying past the target. The selection of these two crucial parameters, possibly combined to polarisation backscattering description might be the key aspect for improving ATR consistency.

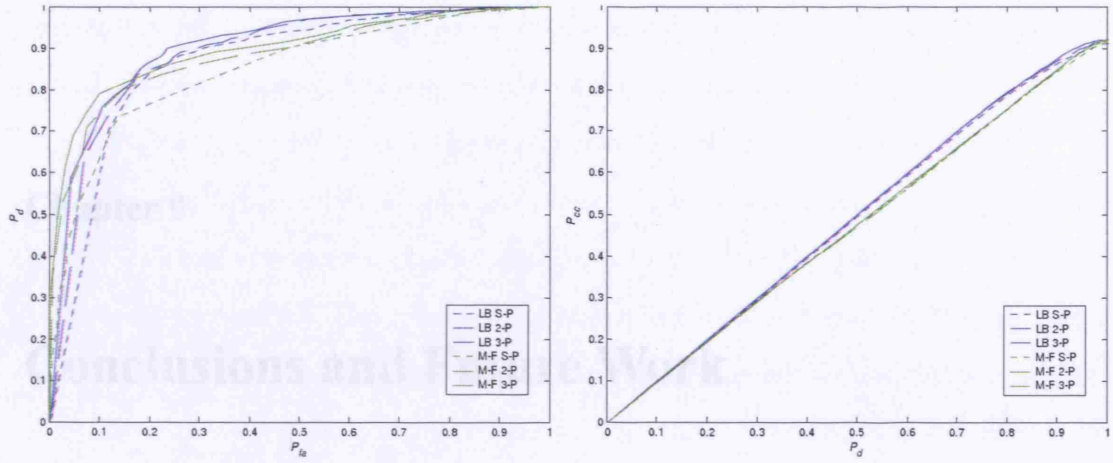

Figure 8.7: *ROC curves for single-perspective (dotted), 2-P (dashed) and 3-P (solid) classifier using large synthesized bandwidth (LB, blue) and eight multiple frequencies (M-F, green).*

8.7 Summary

The classification performance of a Multi-Frequency approach has been compared with the full bandwidth reconstruction methodology and examined using ISAR images from real turntable experiments.

The ROC curves of different RF chirps exhibit variable recognition rates, revealing the importance of the backscattered information dependence on the illuminating frequency. The eventual Multi-Perspective implementation of the classifier suggests the significant reliability of the M-F processing, which is (i) more flexible since a particular combination of centre frequencies or a single one based on the highest ATR response can be ad hoc selected; (ii) it is less demanding in terms of processing and target movement knowledge when compared to the full bandwidth reconstruction of the corresponding wideband signal. Nonetheless, the classification results corroborate the importance of range resolution in terms of classification and the Multi-Perspective effectiveness independence of the resolution used.

Chapter 9

Conclusions and Future Work

The main goal of this thesis was to investigate radar Automatic Target Recognition (ATR) using multiple perspectives sensing a variety of targets. The classical radar target recognition approach based on a single perspective has been demonstrated as being prone to a number of factors that lead to unavoidably poor classification performance. After the overview of the principal radar imaging techniques and the scattering mechanism characterisation of manmade targets, the main sources leading to signature variability and therefore classification performance deterioration have been outlined. Their potential mitigation through the use of multiple aspects has been subsequently introduced.

Data from the Airborne Data Acquisition System (ADAS) and Mobile Instrumented Data Acquisition System (MIDAS) real data from turntable measurements has been processed in order to produce the one- and two-dimensional images of ground vehicle targets. The more populated set of Moving and Stationary Target Acquisition and Recognition (MSTAR) SAR data has also been used to test the Multi-Perspective classifier in order to verify that the results obtained were not data dependent. For this reason, different feature extraction and selection procedures, such as Principal Components Analysis (PCA) and range profile peaks information, have been used as well as a number of pattern recognition approaches implemented (e.g. Bayesian, non-parametric, Artificial Neural Networks, Hidden Markov Model and syntactic classifiers).

The classification performance has been investigated in terms of probability of correct classification (P_{cc}), probability of false alarm (P_{fa}), declaration rate (P_d), generalisation

capabilities (P_{gen}), average reliability (R_{avg}) and Receiver Operator Characteristic (ROC) curves, giving a clear picture of the classifier behaviour. These parameters have been described and formalised as a a metric for evaluating the ATR capabilities in different scenarios.

The benefits brought by multiple perspectives have been analysed firstly implementing a "disjoint" Multi-Perspective classification algorithm (i.e. the two perspectives are treated as independent patterns). The probability of correct classification improvement shows a consistent but non linear improvement with the number of perspectives used. In particular, the capability the classifier to discriminate targets within classes is reduced as the number of processed looks increases, suggesting that a small amount of perspectives is sufficient to achieve a trade-off between system/processing complexity and classification performance improvement over the traditional single-perspective implementation.

These results have also been obtained and analysed at different SNR levels, demonstrating the effectiveness of multiple uncorrelated looks on the correct classification rate. Multiple perspectives therefore augment the accuracy of the main scattering centres description since the lower level scatterers progressively disappear at lower SNR. Nevertheless, the small intensity scatterer returns are fundamental for adequate classification performance.

The probability of correct classification measured for different perspective displacements, using two and three target signatures, shows a consistent classification improvement when the target aspects are sufficiently separated. This is a consequence of the information content link to the cross-correlation properties of the perspectives used. It has been shown that when high correlation occurs, the Multi-Perspective benefits are reduced. Therefore, when the perspectives are closely separated, the classification performance is close to the single perspective case because of the significant degree of cross-correlation. The velocity of perspective decorrelation depends on the resolution and complexity of the target population under consideration. However, it has been found that of the order of five degrees are often sufficient to achieve high Multi-Perspective classification improvements. Moreover, the classification performance is a function of the perspective displacement, and it exhibits different maxima and minima depending on the particular target tested. This is a consequence of the inherent geometrical symmetries of manmade objects.

A crucial feature of radar ATR is the capability of not declaring a class for the "unknown" targets, i.e. the reduction of the false alarm rate. Additional perspectives give relevant cues on the final score to be compared to the fixed rejection threshold.

Multi-Perspective classification can be used also in order to reduce the processing complexity when large apertures are available. Although the full coherent processing yields more accurate scattering centres discrimination and location, therefore leading to higher classification performance, the possible multi-look reconstruction from smaller sub-apertures or their disjoint Multi-Perspective combination still enables significant recognition results. Moreover, they can be used in realtime to update the posterior probability on a cumulative basis, instead of waiting for complete aperture to be synthesised and the full cross resolution to be formed before making a decision.

The use of the perspective displacement information has been investigated when signatures from multiple aspects are attempted to be classified. A number of techniques have been implemented such as the state characterisation using HMM, their hybrid version, the "negative" examples, the target orientation information and, finally, the perspectives concatenation. The use of the topology information, although providing higher classification performance, requires a more accurate target backscattering representation at different aspect angles.

The main difference between disjoint and joint Multi-Perspective radar target classification is the effect of the geometrical symmetries of ground vehicles and, more generally, man-made targets on the classification performance. The capability of recognising a target class at a particular test target orientation is not only a function of its classification score (e.g. its distance) from the corresponding template heading but it is also influenced by other templates at different orientations. The factors leading to signature variability could produce a target signature at orientation α which is close in the feature space to the corresponding library class heading $-\alpha$, where α is taken as the deviation from one of the axes of symmetry for the particular target considered. In a disjoint implementation, this aspect is taken into account since the partial score is the result of comparing the input with the library at any template orientation. Conversely, using the joint implementation, the effect of symmetry aid to classification is reduced by the fact that the angular displacement between perspectives is usually not compatible

with the axis of symmetry. In other words, if the first perspective presents a high score for two templates $\mathbf{t}_{\alpha}$ and $\mathbf{t}_{-\alpha}$ and the second one for $\mathbf{t}_{\alpha'}$ and $\mathbf{t}_{-\alpha'}$, it is unlikely that the perspective displacement $\Delta\varphi_{1,2}$ is such that the symmetry information is exploited.

The average generalisation capability results degraded when the joint Multi-Perspective classification was implemented. On the other hand, the probability of correct classification and the false alarm rate performance are improved by the use of the perspective displacement information. The radar backscattering information is further enhanced when the orientation of the target is estimated and used by the classifier.

If multiple perspectives are available to be processed, and sufficient range resolution can be achieved, the outline of complex targets can be used as source of target geometric features description. The illuminated shape can be extracted and used to provide additional cues of the target and therefore improve classification performance.

The use of multiple signatures originated by illuminating the target at different frequencies has also been investigated, and combined with Multi-Perspective classification. This processing is a valid alternative in terms of classification performance to the synthesis of very high range resolution. Furthermore, when Stepped-Frequency chirp techniques are used, a particular selection of centre frequencies might give similar classification results if compared with the complete set of available chirps, with the results of the possibility of reducing the pulse repetition frequency.

Moreover, similar considerations may apply to acoustic sensors although the different scattering mechanisms characterising sonar backscattering.

The algorithms developed can be incorporated into existing systems in the fields of Air Traffic Control (ATC), Vessel Traffic Management Information Systems (VTMIS) and coastal surveillance systems, navigation radar and remote sensing from air and space. The integration of Earth Observation (EO) and *in-situ* surveillance technologies (Electro-Optics Infrared EO-IR, ground radar, cooperative messaging broadcast such as Automatic Identification System (AIS), etc...) also requires the perspective/technology fusion to obtain reliable and robust classification.

Future Work

Multi-Perspective classification is providing insights into the information content of target backscatter signals. The use of further uncorrelated signatures collected from different aspects allows the improvement of the backscattering description. The methodology described here is a starting point for the investigation of multi-diversity sensor management. A number of parameters can be *ad hoc* varied accordingly to the goal to be achieved.

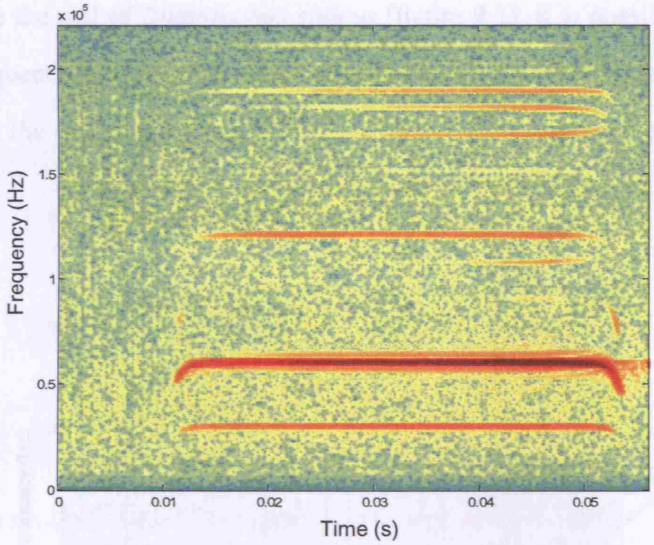

Figure 9.1: *Spectrogram of Rhinolophus sinicus call.*

In Baker et al. (2006) it was shown how echolocating bats continually adapt both their transmit waveform and orientation during the terminal phase of a feeding buzz. In particular they use a combination of spatial and waveform diversity to maximise the information acquired and hence the success of the mission. A detailed analysis shows that the pulse width, PRF and frequency modulation are all being adjusted as the mission develops. In particular, as described in Vespe et al. (paper in revision), the change of pulse length and hyperbolic modulation parameters allows echolocating bats to tune their focus either on range or Doppler information. This concept can be extended to the classification phase. Narrowband calls are used by a wide range of bat species, and the micro-Doppler return given by the flapping wings of preys, produces periodic fluctuations in the Time-Frequency signal representation depending on the

particular trajectory followed.

The micro-Doppler effect is due to micro-motions of parts of a target. In radar ATR, the blades of a helicopter (the most similar micro-motions to a bats prey) show a characteristic spectrum shape around the Doppler shift given by the target velocity (Tait, 2006). It is usually possible to discriminate the number of blades, their velocity and the gear ratio with the rotor and therefore classify the target. This effect might be similarly exploited by echolocating mammals.

If we analyse the call of *Rhinolophus sinicus* (figure 9.1), it is possible to distinguish a long constant-frequency pulse yielding high Doppler resolution two sweeps in frequency at the beginning and the end of the pulse, which are connected to the ranging capability of the bat.

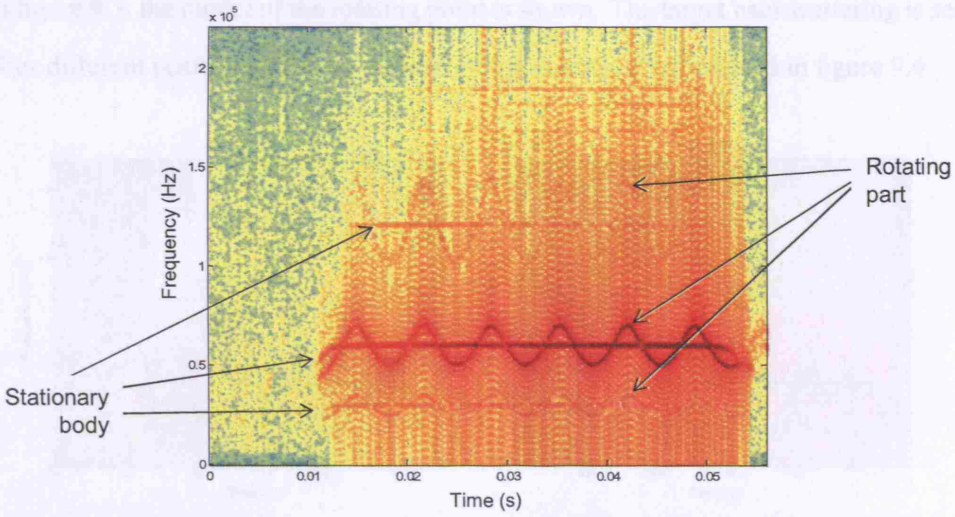

Figure 9.2: *Micro-Doppler return from a simulated rotating point.*

If we simulate the micro-Doppler return given by a rotating object, the resulting spectrogram of the received echo is illustrated in figure 9.2 (Baker et al., 2007). The stationary parts of the target reproduce the same waveform of the transmitted echo, while the rotating scatterer is responsible of the micro-Doppler modulation at different harmonics. The periodicity and the amplitude of this modulation are cues on the wingbeat (periodicity and amplitude), wingspan (amplitude) and therefore dimension of the prey.

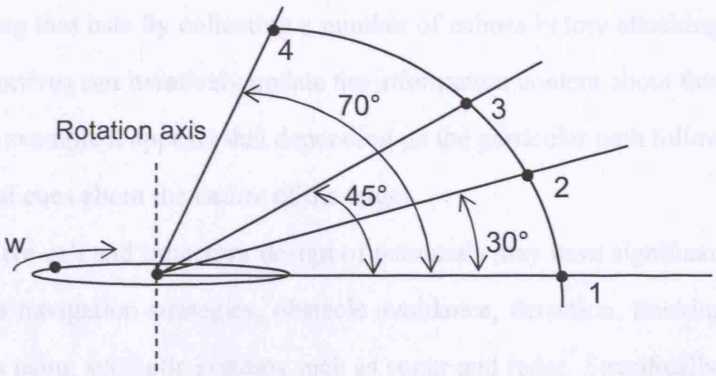

Figure 9.3: *Echoes collected from Different Perspectives. The moving scatterer has a rotation velocity ω and four different perspectives are simulated (identified by the angle with the rotation axis).*

In figure 9.3, the model of the rotating point is shown. The target backscattering is sensed from four different positions, and the corresponding echoes are illustrated in figure 9.4.

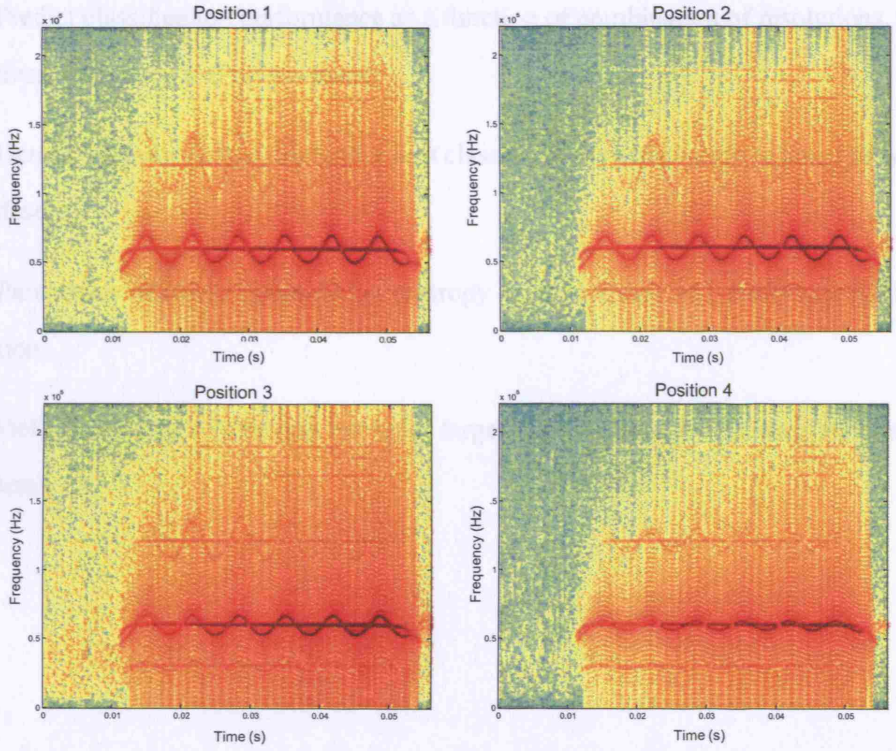

Figure 9.4: *Multi-Perspective micro-Doppler returns from the geometries in figure 9.3.*

Considering that bats fly collecting a number of echoes before attacking, it is clear how multiple perspectives can iteratively update the information content about the prey.

From this example it appears that depending on the particular path followed it is possible to obtain crucial cues about the nature of the target.

The adaptive call and trajectory design of mammals may have significant improvements on autonomous navigation strategies, obstacle avoidance, detection, tracking and classification algorithms using synthetic systems such as sonar and radar. Specifically, optimal sensor management implementations should be scenario dependent in order to iteratively refine the information that could be gleaned.

The main future directions will involve multi-diversity characterisation of radar backscattering and can be therefore summarised as follows:

- Establish a Linkage between radar parameters, target signature information and classification using multi-diversity;
- Predict classification performance as a function of combination of resolutions, perspectives, frequency and polarisation;
- Determine radar design drivers for best classification performance in terms of waveform design;
- Parameters adaptive optimisation (entropy minimisation) and multi-source classification;
- Deliver algorithmic designs for radar target classification for existing and future systems.

Appendices

Appendix A

Naïve Bayesian Classifier

Naïve Bayesian classifier is a statistical method for pattern recognition. It is called on to reduce the decision-making problem to simple calculations of feature probabilities. The training set is stored in different libraries of templates, thus, the number of classes is known as well as sets of their representative vectors. Therefore, the algorithm can be considered characterised by supervised strategy.

Naïve Bayesian classifier is based on Bayes theorem to calculate the posterior probability of classes conditioned on the given unknown feature vector. Two fundamental probability definitions are listed below:

- *Prior Probability* $P(E)$ also known as unconditional probability, of an event E is the probability assigned to an event in absence of knowledge supporting its occurrence or absence.

- *Posterior Probability* $P(E|e)$, often called conditional probability, of an event E conditioned to an evidence e is the probability of an event given some evidence.

Considering a feature vector $X = x_1, ..., x_n$ whose class is unknown and C_i the hypotheses that it belongs to class $i = 1, ..., n_c$, the posterior probability $P(C_i|X)$ of X conditioned to C_i is the probability that X belongs to that class. Similarly, $P(X|C_i)$ is the probability that any sample from class i have the predictor attributes of X provided that X is in class i. This is often referred to as *Likelihood*. The prior probability $P(C_i)$ is independent of X, $P(X)$ and $P(X|C_i)$ and it can be estimated from available data. The posterior probability $P(C_i|X)$ is

not known but can be estimated by applying the Bayes theorem through the calculation of the other probabilities:

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{\sum_{h=1}^{n_c} P(X|C_h)P(C_h)} \tag{A.1}$$

The elements on the right hand side of equation A.1 can be easily calculated from known variables and it is possible to implement the final decision on the class. The Naïve Bayesian classifier is based on the assumption of the class conditional independence of each element of the feature vector X. The algorithm can be divided into five stages as follows:

1. Each object from the population P is represented by an n-dimensional vector X, forming a sequence of n attribute values $a_1, a_2, ..., a_n$.

2. Given an unknown feature vector X, the classifier will decide that X belongs to the class which has the highest posterior probability. That means the Naïve Bayesian classifier assigns X to the class i if and only if

$$P(C_i|X) > P(C_j|X) \quad \text{for} \quad 1 \leq j \leq n_c,\ j \neq i \tag{A.2}$$

   By applying Bayes theorem, the maximum posteriori hypothesis can be calculated using equation A.1.

3. Since prior probability $P(X)$ remains constant for all classes, only the term $P(X|C_i)P(C_i)$ needs to be maximised. If the class prior probabilities are not known, then they are commonly assumed to be equally probable, that means $P(C_1) = P(C_2) = ... = P(C_{n_c})$. Another criterion to estimate class prior probabilities is in relation to the number of template vectors contained in each training set K_i:

$$P(C_i) = \frac{K_i}{\sum_{j=1}^{n_c} K_j} \quad \text{for} \quad 1 \leq j \leq n_c \tag{A.3}$$

4. The "naïve" assumption is now applied in order to reduce the computational effort to calculate $P(X|C_i)$ which is a consequence of high dimensional feature vectors X. Therefore, statistical independence of attributes of feature vectors is assumed, that is the

values of the attributes are conditional independent of each other. Hence the posterior probabilities for each class can be approximated as follows:

$$P(X|C_i) = \prod_{l=1}^{n} P(x_l|C_i) \tag{A.4}$$

When calculating each $P(x_l|C_i)$, if the attributes belong to a discrete set of values, then the algorithm implements

$$P(x_l|C_i) = \frac{K_{x_l}}{K_i} \tag{A.5}$$

where K_{x_l} is the number of training samples of class i having x_l equal to the particular attribute a_l, and K_i is the number of training samples of class i. Conversely, if values a_l are continues valued, then the attribute is assumed to have Gaussian distribution and each conditional probability is calculated as:

$$P(x_l|C_i) = \frac{1}{\sigma_{C_i}\sqrt{2\pi}} e^{-\frac{(x_l-\mu_{C_i})^2}{2\sigma_{C_i}}} \tag{A.6}$$

where μ_{C_i} and σ_{C_i} are the mean value and the standard deviation respectively of the element x_l for the training samples of class i.

5. Following this procedure, the Naïve Bayesian classifier can perform the decision for the feature vector presented as an input by calculating the likelihood ratio test:

$$\text{if} \quad \frac{P(X|C_i)}{P(X|C_j)} > \frac{P(C_j)}{P(C_i)} \quad \Rightarrow \quad X \in \text{class } i \tag{A.7}$$

The main advantage of the Naïve Bayesian classifier is that the probability of error $P_E(X)$ can be measured for each decision between two classes a and b:

$$P_E(X) = \min\Big\{P(X|C_a)P(C_a), P(X|C_b)P(C_b)\Big\} \tag{A.8}$$

However, a first drawback of this method is the number of assumptions required and certain *a priori* information is needed, i.e. the *a priori* distributions $P(C_i)$ over the classes, as well as the conditional probabilities $P(X|C_i)$, must be estimated which can be a difficult task if suitable distributions are not assumed. Even if these assumption are reasonable, the most crucial disadvantage of statistical approaches to classification is the fact they have essentially linear results and it is a remarkable defect for many difficult classification problems.

Appendix B

K-Nearest Neighbour Classifier

A non parametric method for classification is K-Nearest Neighbour (K-NN) classifier. The approach consists of measuring and minimising the number of K distances from the object to the elements of the training set. Considering an n-dimensional feature vector $X = (x_1, x_2, ..., x_n)$ whose class is unknown and a number n_c of classes C_i formed by n_i template vectors each, the K-NN algorithm can be divided into three steps:

1. After defining the number K, the distance $d_{t,i}$ from X to each representative vectors $T_{t,i}$ of class C_i are computed and stored

$$d_{t,i} = d(X, T_{t,i}) = \|X - T_{t,i}\| \tag{B.1}$$

   for each template index $t = 1, ..., n_i$ and class index $i = 1, ..., n_c$.

2. The K minimum values of $d_{t,i}$ are selected from each class forming n_c K-dimensional vectors D_i in ascending order.

3. The elements of vectors D_i are compared, selecting the smallest values. Each selected value forms a vote for the class it belongs to. Ultimately, the classifier decides for the class with greatest number of votes over the K obtained.

As can be observed from figure B.1, the choice of the number K is a determinant factor for K-NN classifier decision: if $K = 1$, the nearest region identified by the smallest dotted circle suggests the outcome of the classifier is class "1" in view of the fact that the nearest template vector to the novel X belongs to class "1". If $K = 2$ the K-NN outcome is unknown

and, if $K = 3$ the decision is for class "2". Generally, for a given classification problem, a small K leads to a large misclassifications. On the other hand, setting K to large values implies high computational efforts and not necessarily a better classification performance. A solution is to classify known patterns (*validation set*) with different K and establish the optimum K for the specific classification problem *(Cross-Validation)*.

The advantage of this decision-making procedure is simplicity and the non-linear mapping of the input. The disadvantages are the difficulty to establish the optimum K and the fact that K-NN ignores those clusters having a central region where the feature vectors are more frequently represented than outside of this region.

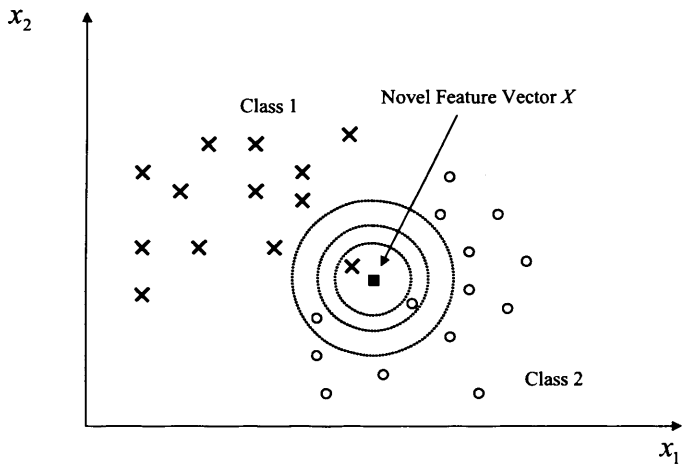

Figure B.1: *K-nearest neighbors classifier: different associations results with respect to different K values.*

Appendix C

Feed-forward Artificial Neural Networks

Artificial Neural Networks (ANNs) arise from the idea of reproducing some of the abilities and functionalities of the human brain. The principal ANN field of application is pattern recognition: given a feature vector, the network is able to analyse it and associate an output corresponding to a particular class of objects.

ANN is fundamentally an adaptive system that learns how to execute a particular task through examples. The information contained in the training samples is used to set internal parameters of the network. The recognition process, when learning methods are employed, is also called *operational mode*: the system maps the input feature vector into the output number of the class. If the system also classifies (that means to organise and divide the population into sub-population sets), then it has a *learning mode*. There are two categories of machine learning:

- *Supervised Training*: during the learning phase a "supervisor" establishes how the classifier is efficient and subsequently corrects the decision parameters in order to optimise recognition. Each element of the training set is given to the classifier as well as the wanted output.

- *Self Organisation (Unsupervised training)*: the system performs both training and operational processes on any input vector. The classifier, in an autonomous way, looks for peculiar properties of data at the input and it learns how to link them to the output.

In either the supervised or unsupervised training mode, the system finds a representative vector

for each class and, during the operational model, it compares the input vector to each template vector for each class. A decision is made by the recogniser by measuring the greatest similarity between them.

In this work, *Feed-Forward* ANNs supervised with *Back-propagation* strategy have been investigated and implemented. A neural network consists of a number of fundamental units called *neurons*, able to communicate to each other using connections called *synapsis*. Neuron connections can be organised in different ways according to the particular network that has to be created. Starting from a given example, the network processes by successive steps, its own internal representation of data yielding the final output.

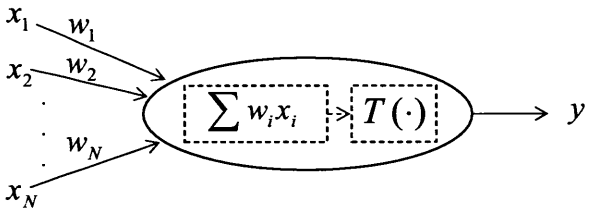

Figure C.1: *Graphic representation of a generic neuron, the elementary ANN unit.*

Each neuron is an independent unit having multiple input connections $(x_1, x_2, ..., x_n)$ and a single output port y. As depicted in figure C.1, each input connection has a different importance defined by a quantity w_i called *weight*. The neuron sums up all the inputs taking into account their connection weights. Eventually, it calculates the output by applying a *transfer function* T. Signals either in the input or output can assume continuous values in the range [0,1]. Weights can be either positive or negative. Moreover, every neuron has a bias weight w_b in order to regulate the neuron output. *Internal activation* γ of a neuron is defined as the weighted sum of all the input signals:

$$\gamma = x_1 w_1 + x_2 w_2 + ... + x_n w_n \tag{C.1}$$

The output signal y is known as *activity* and it is the result of the transfer function T which maps a sum γ into a proper range of output values. A number of activation functions can be chosen, as shown in figure C.2. In this work, a *unipolar sigmoid* (or *logistic*) function

has been used, whose expression is:

$$y = \frac{1}{1 + \alpha e^{-(\gamma - a)}} \tag{C.2}$$

where α is the decay rate and a is the bias that shifts the function centre.

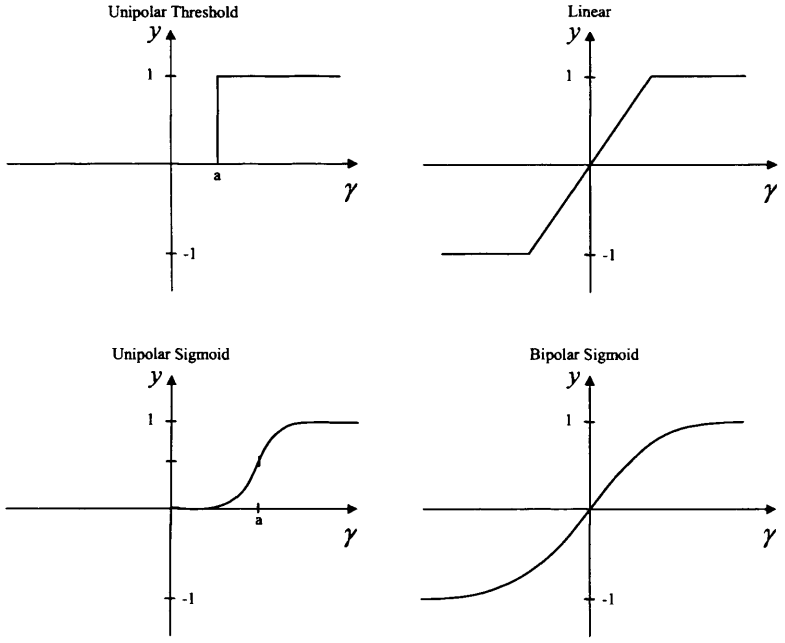

Figure C.2: *Typical Transfer Functions.*

It has been demonstrated by Bridle (1990), "when used for classification, MLPs can generate good estimates of posterior probabilities of the output classes conditioned on the input pattern." It is possible to interpret the output of ANN as a posterior probability. As a consequence of the sigmoid function, the output layer always gives a response in the range between zero and one. Nevertheless, the sum of the output vector is not required to be one in order to perform the learning phase, whereas in order to give an estimate of the class posterior probability the sum has to be one. This can be done by using the *softmax* (Bourlard and Morgan (1994)) activation function on the output layer, defined as follows:

$$y_i = \frac{\exp(g_i)}{\sum_{l=1}^{C} \exp(g_i)} \tag{C.3}$$

where g_i is the sum of the weighted input vector to the i-th output neuron y_i and C the dimension of the output vector, i.e. the number of classes of the population set.

According to the weight values, by varying the input vector, the neuron produces a different mathematical function for the activation of y. In order to produce complex functions, neural networks are made up different *layers*. The first layer is the input X. This layer is not involved in the global ANN computation in view of the fact that it has no weights in its connections. After the input, a variable number of intermediate layers known as *hidden layers* follow. With proportion to the problem complexity and the mathematical function which approximate the mapping, a single hidden layer contains a variable number of neurons. Using two hidden layer, it is possible to reduce the global number of neurons. The final hidden layer is ultimately connected to the *output layer*, which provides the solution to the network.

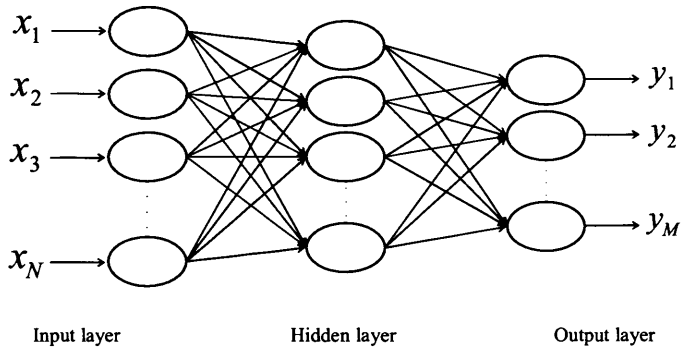

Figure C.3: *Single hidden layer Feed-forward Neural Network.*

In figure C.3, a single-layer Feed-forward neural network is represented. The term Feed-forward is related to the single forward direction of data flow in the network. Nevertheless, as explained in Appendix A, a back-propagation learning strategy allows a feedback of information during the training mode.

As previously discussed, a classifier needs to learn how to make decisions. The learning phase of ANNs may be executed by a number of different approaches and algorithms with the same aim of creating an internal representation of data. The two main categories have previously been discussed: *supervised* and *unsupervised*. In the unsupervised strategy, the learning algorithm is based on the history of the system and it trains weights in order to reach

some internal cost function. Unsupervised networks are used when data are not either well organised or classified and the reference of different classes is not known. In these cases the ANN is also known as *self-organising* because it finds in an autonomous way the classification criteria. Supervised learning strategy is adopted for those problems characterised by well defined results. With this approach, the network is able to generalise the classification from examples given as input. As shown in figure C.4, during the first learning cycle (*epoch*), two vectors are given to the network: the input and the requested output. The ANN calculates the output which could be either correct or incorrect. Subsequently, the supervisor corrects the weights of each single connection at each epoch, until an error is minimised. This procedure is repeated for each input-output pair.

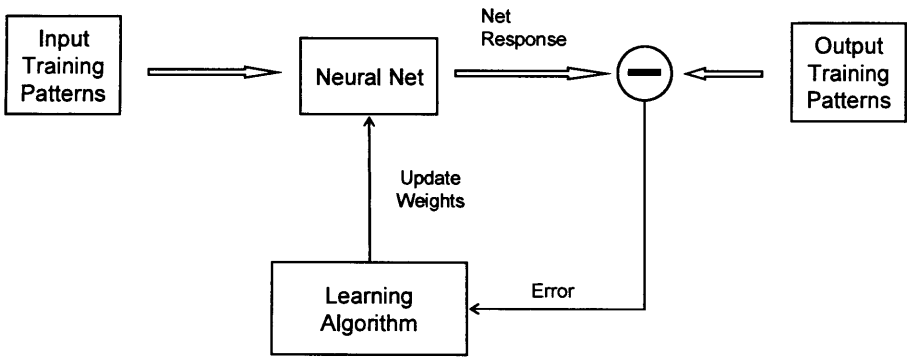

Figure C.4: *Supervised learning strategy for FANNs with back-propagation.*

This particular learning phase is also known as back-propagation of the error. As defined by Arbib (1995):

1. Back-propagation is a procedure for efficiently calculating the derivatives of some output quantity of a nonlinear system, with respect to all inputs and parameters of that system, through calculations proceeding backwards from outputs to inputs. It permits implementation on parallel hardware.

2. Back-propagation is any technique for adapting the weights of parameters of a nonlinear system by somehow using such derivatives or their equivalent.

Weights are set to small random values and all the template vector whose class is known

are put as inputs to the network. A procedure to correct the weights is applied until a certain distance between the actual and expected outputs is achieved. In order to implement this concept, the *Generalised Delta Rule* (GDR) is used.

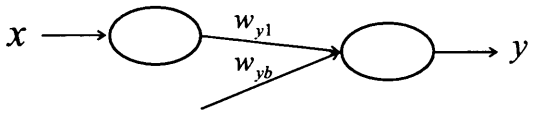

Figure C.5: *Single input-output network.*

Considering figure C.5, the neuron y is the single output of the network, x the single input, and w_{y1} and w_{yb} respectively the output weight linked to the input and the output bias weight. In order to find the error E_y, the difference between the output and the desired output value d is calculated: $E_y = y - d$. The delta of the y neuron δy is then calculated by multiplying the error times the derivative of output transfer function T:

$$\delta y = E_y \cdot \frac{\partial T(\gamma_y)}{\partial \gamma_y} \tag{C.4}$$

where γ is the activation of the output neuron: $\gamma_y = x w_{y1} + w_{yb}$. This is done so that the delta represents the velocity of error variation with respect to internal activation variations. Each single weight at epoch $e + 1$ is then modified as follows:

$$w_{y1}^{e+1} = w_{y1}^{e} + (\epsilon\, \delta y\, x) \tag{C.5}$$

$$w_{yb}^{e+1} = w_{yb}^{e} + (\epsilon\, \delta y) \tag{C.6}$$

Equations C.5 and C.6 show that each weight is modified by a quantity equal to the neuron delta times the signal present in the synapsis it represents. It is ultimately multiplied by the *learning rate* ϵ of the network which is *a priori* chosen in the range interval (0,1). The greater the learning rate, the faster the variation of weight values during the training phase. On the other hand, if a large learning rate is set, large errors and oscillations can result. Viceversa, if the learning rate is close to zero, the time required for training the network remarkably increases and greater accuracy is not always achieved: it is possible to stop the learning in a local minimum of the error function. If the neuron belongs to the hidden layer (figure C.6), it

is not possible to evaluate its error as a distance between expected output and the actual one: in this case the previous error E is back-propagated to the hidden layer h (from here the name of the method).

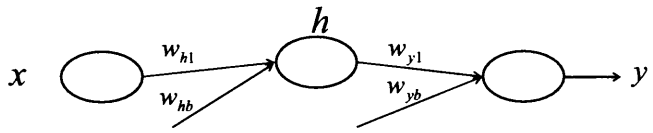

Figure C.6: *Single input-output network with a hidden layer.*

Therefore, the error for the hidden layer is $E_h = \delta y \cdot w_{y1}$. As a result, the delta value for the hidden layer can be expressed as:

$$\delta h = E_h \cdot \frac{\partial T(\gamma_h)}{\partial \gamma_h} \tag{C.7}$$

After calculating both error and delta values, it is possible to change the weights for the hidden layer in the same way as the output layer:

$$w_{h1}^{e+1} = w_{h1}^{e} + (\epsilon\ \delta h\ h) \tag{C.8}$$

$$w_{hb}^{e+1} = w_{hb}^{e} + (\epsilon\ \delta h) \tag{C.9}$$

These results can be extended to multi-dimensional input and output artificial neural networks. A global procedure for training a FANNs is now summarised.

1. Weights of hidden and output layers are set to small random values.
2. For each pair of inputs and expected output vectors (example):
   (a) the input vector is given to the network, is executed and an output vector is calculated;
   (b) errors and delta values are determined for each neuron;
   (c) weights are modified.
3. If the maximum value of any of the network output neurons, during any of the examples, is greater than the maximum allowed, go to point 2.

The execution mode, that is the calculation of the output vector $Y = (y_1, y_2, ..., y_M)$, of neural network is performed just by multiplying the input vector $X = (x_1, x_2, ..., x_N)$ by the weights of the hidden layers $\{w_{h,nj} : n = 1, ...N; j = 1, ..., J\}$ and then by the output layer's ones $\{w_{y,jm} : j = 1, ...J; m = 1, ..., M\}$. This is one of the advantages of neural networks. As a consequence, the speed of the execution mode for neural networks is one of their principal advantages. Furthermore, a database of template vectors is not required after training the network, so the classifier can be seen as a number of matrices which might be easily transmitted from nodes of a network of classifiers. Nevertheless, one of the most interesting advantages of artificial neural networks is the possibility to map the input vector into the output one also for problems with no linear separation of the classes. On the other hand, the principal drawback of neural networks is the difficulties in determining the right learning rate and the possibility of local minima error conversion during the learning mode.

Appendix D

Hidden Markov Model (HMM)

Hidden Markov Models (HMM) were first applied in speech recognition in the early 1970-s (Rabiner, 1989). A HMM is a statistical model of a system assumed to be a process with unknown parameters. The hidden parameters are evaluated from the observable ones and the obtained model can ultimately be used to perform further analysis, for example for pattern recognition applications.

In Markov models, the state S is visible to the observer, and therefore the state transition probabilities are the only parameters. When the Markov model is "hidden", the state sequence is not visible to an external observer.

A HMM is identified by the notation $\lambda(A, B, \pi)$ and a set of N states, where the parameters are defined as follows:

- In case of first order HMM, $A = a_{i,j}$ is the set of state transition probabilities to the current state q_t from the previous one q_{t-1}:

$$a_{i,j} = P(q_t = i | q_{t-1} = j), \quad 1 \leq i, j \leq N \tag{D.1}$$

  where the indices i and j are in the discrete range $[1, N]$. The transition probabilities follow the stochastic properties:

$$\sum_{j=1}^{N} a_{i,j} = 1 \quad a_{i,j} \geq 0 \quad \forall i, j \tag{D.2}$$

- The probability distribution for each state $B = \{b_i(k)\}$, also known as *emission proba-*

bility,:

$$b_i(k) = P(o_t = v_k | q_t = i), \quad 1 \leq k \leq M, \quad 1 \leq i \leq N \tag{D.3}$$

where v_k is the k-th of the M symbols alphabet and o_t is the observation at time t. The state probabilities satisfy the stochastic constrains:

$$\sum_{k=1}^{M} b_i(k) = 1 \quad b_i(k) \geq 0 \quad \forall i, k \tag{D.4}$$

- The set of initial state distribution $\pi = \pi_i$ where the probability of starting the sequence in the i-th state is:

$$\pi_i = P(q_1 = i), \quad 1 \leq i \leq N \tag{D.5}$$

We now concentrate on the particular case where a sequence $O = o_1, o_2, ..., o_l$ is observed and the probability that the model λ generated it with a probability $P(O|\lambda)$. If the observation is quantised into a set of templates, then the HMM is *discrete*. Otherwise, if quantisation of the observation is not performed, the HMM is said *continuous*. The likelihood of observing O for a particular state sequence $Q = q_1, q_2, ..., q_l$ can be expressed as follows:

$$P(O|Q, \lambda) = b_{q_1}(o_1) b_{q_2}(o_2) ... b_{q_l}(o_l) \tag{D.6}$$

The probability of a particular state sequence Q generated by the model λ is represented by the following equation:

$$P(Q|\lambda) = \pi_1 a_{q_1,q_2} a_{q_2,q_3} ... a_{q_{l-1},q_l} \tag{D.7}$$

As a consequence, the joint probability of a set of observation O and a state sequence Q, given the model λ is:

$$P(O, Q|\lambda) = P(O|Q, \lambda) P(Q|\lambda) \tag{D.8}$$

The probability of a HMM to generate a set of observed patterns, hiding the particular state sequence is therefore measured and summed over all possible state sequences:

$$P(O|\lambda) = \sum_{\forall Q} P(O, Q|\lambda) = \sum_{\forall Q} P(O|Q, \lambda) P(Q|\lambda) \tag{D.9}$$

We here limit the description of the calculation of the probability $P(O|\lambda)$ to the *forward-backward* procedure introduced by Baum and Egon (1967) as the basis for the application discussed in Chapter 6. The variable $\alpha_t(i)$ represents the probability that the sequence of observations originated by the model λ presents the current state at time t is S_i:

$$\alpha_t(i) = P(O, q_t = S_i|\lambda) \tag{D.10}$$

Inductively, $\alpha_t(i)$ is initialised for $t = 1$:

$$\alpha_1(i) = \pi_i b_i(o_1) \tag{D.11}$$

This is valid for $1 \leq i \leq N$. The forward variable α having S_i as final state is subsequently calculated at time $t + 1$:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] b_j(o_{t+1}) \tag{D.12}$$

This is calculated for $1 \leq t \leq T$. Eventually, the resulting probability that given a model λ, the sequence O is observed is calculated as the sum of the forward variable for any possible final state in the set:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{D.13}$$

If we consider a simple "Left-Right" HMM in figure (D.1) formed by three states $S = \{S_1, S_2, S_3\}$.

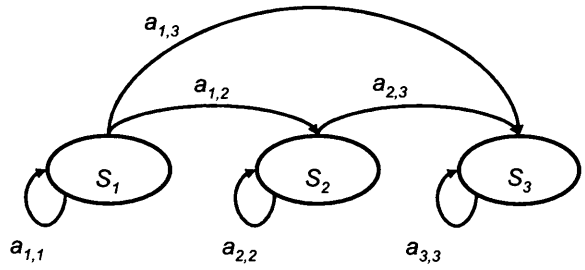

Figure D.1: *Three-state, "Left-Right" Hidden Markov Model (HMM).*

Assuming a sequence of two observations $O' = o_1, o_2$, the corresponding transition prob-

abilities, probability state distribution and initial state distributions are:

$$A = \{a_{i,j}\} = \begin{pmatrix} a_{1,2} & a_{1,3} & a_{1,3} \\ 0 & a_{2,2} & a_{2,3} \\ 0 & 0 & a_{3,3} \end{pmatrix} \tag{D.14}$$

$$B = \{b_i(k)\} = \begin{pmatrix} b_1(o_1) & b_1(o_2) \\ b_2(o_1) & b_2(o_2) \\ b_3(o_1) & b_3(o_2) \end{pmatrix} \tag{D.15}$$

$$\pi = \{\pi_1, \pi_2, \pi_3\} \tag{D.16}$$

Then using equations D.11, D.12 and D.13, the probability $P(O'|\lambda)$ can be calculated:

$$\begin{aligned} P(O'|\lambda) = {} & [\pi_1 b_1(o_1) a_{1,1}]\, b_1(o_2) + [\pi_1 b_1(o_1) a_{1,2} + \pi_2 b_2(o_1) a_{2,2}]\, b_2(o_2) + \\ & + [\pi_1 b_1(o_1) a_{1,3} + \pi_2 b_2(o_1) a_{2,3} + \pi_3 b_3(o_1) a_{3,3}]\, b_3(o_2) \end{aligned} \tag{D.17}$$

The Left-Right HMM is suitable to model Multi-Perspective states where the perspectives are collected from the same projection plane since the angular displacement is a scalar. If this approximation is not valid (i.e. the perspective displacement between two consecutive looks is a two-dimensional vector having azimuth and elevation components) this model does no longer hold.

Appendix E

Equivalence of Joint and Disjoint Naïve Bayesian Classifier

The "naïve" assumption of elements statistical independence results in the ineffective use of the perspective topology Ψ. If we suppose to collect two perspectives X_1 and $X2$, and a two-class population set, the M-P disjoint classifier computes the four likelihoods:

$$P(X_1|C_1) = a \quad P(X_1|C_2) = b \quad P(X_2|C_1) = c \quad P(X_2|C_2) = d \tag{E.1}$$

The parallel calculation of the single perspective stages leads to the computation of the posterior probabilities:

$$P(C_1|X_1) = \frac{P(X_1|C_1)}{P(X_1|C_1) + P(X_1|C_2)} = \frac{a}{a+b} \tag{E.2}$$

$$P(C_2|X_1) = \frac{P(X_1|C_2)}{P(X_1|C_1) + P(X_1|C_2)} = \frac{b}{a+b} \tag{E.3}$$

$$P(C_1|X_2) = \frac{P(X_2|C_1)}{P(X_2|C_1) + P(X_2|C_2)} = \frac{c}{c+d} \tag{E.4}$$

$$P(C_2|X_2) = \frac{P(X_2|C_2)}{P(X_2|C_1) + P(X_2|C_2)} = \frac{d}{c+d} \tag{E.5}$$

where we assumed the prior probabilities $P(C_i)$ to be uniformly distributed over the two classes. The disjoint processing described in Chapter 5 yields the computation of the posterior probabilities:

$$P(C_1|X_1, X_2) = \frac{P(C_1|X_1)P(C_1|X_2)}{P(C_1|X_1)P(C_1|X_2) + P(C_2|X_1)P(C_2|X_2)} = \frac{ac}{ac+bd} \tag{E.6}$$

$$P(C_2|X_1,X_2)=\frac{P(C_2|X_1)P(C_2|X_2)}{P(C_1|X_1)P(C_1|X_2)+P(C_2|X_1)P(C_2|X_2)}=\frac{bd}{ac+bd} \tag{E.7}$$

If we now consider the joint approach, the new class templates obtained by concatenating the reference vectors with the corresponding displaced by the angle Φ are C_1' and C_2'. The combined feature is vector $X_\Phi = [X_1; X_2]$ and the likelihood to be generated by class i is evaluated using equation A.4 and A.6. The new template library for the i-th class, is formed by the concatenation of the τ n-dimensional vectors:

$$T_i'=\begin{bmatrix} \mathbf{t}_1^k & \mathbf{t}_2^i & \dots & \mathbf{t}_{\tau_i}^k \\ \mathbf{t}_{1+\Phi}^k & \mathbf{t}_{2+\Phi}^i & \dots & \mathbf{t}_{\tau_i+\Phi}^k \end{bmatrix}=\begin{bmatrix} \mathbf{t}_1(1) & \mathbf{t}_2(1) & \dots & \mathbf{t}_\tau(1) \\ \vdots & & \ddots & \\ \mathbf{t}_1(n) & \mathbf{t}_2(n) & \dots & \mathbf{t}_\tau(n) \\ \mathbf{t}_{1+\Phi}(1) & \mathbf{t}_{2+\Phi}(1) & \dots & \mathbf{t}_{\tau+\Phi}(1) \\ \vdots & & \ddots & \\ \mathbf{t}_{1+\Phi}(n) & \mathbf{t}_{2+\Phi}(n) & \dots & \mathbf{t}_{\tau+\Phi}(n) \end{bmatrix} \tag{E.8}$$

Since the class parameters are calculated over all possible states, the new template set is represented by the mean values μ_i and standard deviation σ_i.

$$T_i' \Longrightarrow \begin{bmatrix} \mu_1, & \sigma_1 \\ \vdots & \\ \mu_n, & \sigma_n \\ \mu_1, & \sigma_1 \\ \vdots & \\ \mu_n, & \sigma_n \end{bmatrix} \tag{E.9}$$

As a consequence, when the joint likelihood is calculated, the topology information is lost:

$$P(X_\Phi|C_1')=P(X_1|C_1)P(X_2|C_1)=ac \tag{E.10}$$

$$P(X_\Phi|C_2')=P(X_1|C_2)P(X_2|C_2)=bd \tag{E.11}$$

and the final posterior probabilities to be maximised:

$$P(C_1'|X_\Phi) = \frac{ac}{ac + bd} \tag{E.12}$$

$$P(C_2'|X_\Phi) = \frac{bd}{ac + bd} \tag{E.13}$$

This is equal to the disjoint results (equations E.6 and E.7). In conclusion, the "naïve" Bayesian classifier is not suitable for a joint M-P implementation and other methodologies need to be considered in order to improve the classification performance with respect to the disjoint version.

Bibliography

ANSI/IEEE Standard Dictionary of Electrical and Electronics Terms. ANSI/IEEE std 100-1984, New York, IEEE press, 1984.

A. Aprile, A. Mauri, and D. Pastina. Real time rota-tional motion compensation algorithm for focusing spot-SAR/ISAR images in case of variable rotation rate. In *Proc. of the EUropean RADar (EURAD) Conference*, pages 141–144, Amsterdam, October 2004.

M. A. Arbib. *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA, 1995.

ASCIET. All Service Combat Identification Evaluation Team, http://www.xs4all.nl/~aerocom/asciet/, 2000.

D. A. Ausherman, A. Kozma, J. L. Walker, H. M. Jones, and E. C. Poggio. Developments in radar imaging. *IEEE Transaction on Aerospace and Electronic Systems*, AES–20:363–399, July 1984.

C. J. Baker and A. L. Hume. Netted radar sensing. *IEEE Aerospace and Electronic Systems Magazine*, 18:3–6, 2003.

C. J. Baker and M. Vespe. Multi-perspective imaging and image interpretation. In *NATO ASI Imaging for Detection and Identification*. ASI NATO, Springer, August 2006.

C. J. Baker, H. D. Griffiths, G. Jones, and M. Vespe. 50 million years of waveform diversity. In *Proceedings of the IET conference on waveform diversity and design*, London, November 2006.

C. J. Baker, M. Vespe, and G. Jones. Diversity strategies: Lessons from natural systems. In *Paper submitted to the International Waveform Diversity and Design Conference*, Pisa, 2007.

M. Baker and T.E. Boult. Pruning bayesian networks for efficient computation. In *In Proc. Sixth Conf. on Uncertainty in Artificial Intelligence*, pages 257–264, Cambridge, 1990.

L. E. Baum and J. A. Egon. An inequality with applications to statistical estimation for probabilistic functions of a Markov process and to a model for ecology. *Bulletin of the American Meteorological Society*, 73:360–363, 1967.

M. R. Bell and R. A. Grubbs. JEM modeling and measurement for radar target identification. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):73–87, January 1993.

G. R. Benitz. High-definition vector imaging for synthetic aperture radar. In *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, October 1997.

B. Bhanu and G. Jones. Recognizing target variants and articulations in synthetic aperture radar images. *Optical Engineering*, 39(3):712–723, March 2000.

G. Bienvenu and L. Kopp. Adaptivity to background noise spatial coherence for high resolution passive methods. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 307–310, April 1980.

C. M. Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, 1995.

E. C. Botha, E. Barnard, and C.J. Barnard. Feature-based classification of aerospace radar targets using neural networks. In *Neural Networks*, volume 9, pages 129–142, 1996.

H. Bourlard and N. Morgan. *Connectionist Speech Recognition - A Hybrid Approach.* Kluwer Academic, 1994.

J. S. Bridle. Probabilistic interpretation of feedforward classification network outputs with relationships to statistical pattern recognition. In *Neurocomputing - Algorithms, Architectures and Applications*, volume 6, pages 227–236, Berlin, 1990. Springer-Verlag.

Hung-Chih Chiang, R.L. Moses, and L.C.; Potter. Model-based classification of radar images. *IEEE Transactions on Information Theory*, 46:1842–1854, 2000.

M. N. Cohen. An overview of radar-based, automatic, noncooperative target recognition techniques. In *IEEE Int. Conference on Systems Engineering*, pages 29–34, August 1991.

M. N. Cohen. Radar and multi-sensor target recognition. In *2006 IEEE Radar Conference Tutorial*, Verona (NY), April 2006.

J. C. Curlander and R. N. McDonough. *Synthetic aperture Radar: Systems and Processing.* Wiley, New York, 1991.

L. J. Cutrona. *Radar Handbook, M.I. Skolnik*, chapter Synthetic Aperture Radar. McGraw-Hill, 1970.

M. D. Desai and W. K. Jenkins. Convolutional backprojection image reconstruction for spotlight-mode synthetic aperture radar. *IEEE Transaction on Image Processing*, 1(4):505–517, October 1992.

R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification.* John Wiley and Sons, 2001.

J. P. Egan. *Signal Detection Theory and ROC Analysis.* Academic Press, New York, 1975.

Defence Science Board Task Force. Report on combat identification, www.acq.osd.mil/dsb/reports/combatidentification.pdf, 1996.

W. Genyuan, X. Xiang-Gen, and V. C. Chen. Three-dimensional ISAR imaging of maneuvering targets using three receivers. *IEEE Transaction on Image Processing*, 10(3):436–447, March 2001.

M.J. Gerry, L. C. Potter, I. J. Gupta, and A. Van Der Merwe. A parametric model for synthetic aperture radar measurements. *IEEE Transactions on Antennas and Propagation*, 47(7): 1179–1188, July 1999.

D. Gross. A neural network ATR for high range resolution radar signature recognition of moving ground targets. In *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, volume 2, pages 1235–1239, October 1999.

Li Hai-bin, Zhang Yun-hua, and Wu Jie;. Sidelobes and grating lobes reduction of stepped-frequency chirp signal. In *IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, volume 2, pages 1210–1213, August 2005.

Rong Hu and Zhaoda Zhu. Researches on radar target classification based on high resolution range profiles. In *Proceedings of the IEEE Aerospace and Electronics Conference*, volume 2, pages 951–955, 1997.

Xiao Huaitie, Zhuang Zhaowen, Chen Zhenpin, He Songhua, and Guo Biao. Aircraft target recognition using adaptive time-delay neural network. In *Proceedings of the Aerospace and Electronics Conference*, volume 2, pages 764–768, 1997.

S. Hudson and D. Psaltis. Correlation filters for aircraft identification from radar range profiles. *IEEE Transactions on Aerospace and Electronics Systems*, 29(3):741–748, July 1993.

K. Ikeuchi, M. D. Wheeler, T. Yamazaki, and T. Shakunaga. SAR target recognition based on invariant histograms and deformable template matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco*, 1996.

S. P. Jacobs and J. A. O'Sullivan. Automatic target recognition using sequences of high resolution radar range-profiles. *IEEE Transactions on Aerospace and Electronic Systems,*, 36 (2):364–381, April 2000.

A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.

D. Jalihal and L.W. Nolte. Signal detection theory and reconstruction algorithms-performance for images in noise. *IEEE Transactions on Biomedical Engineering*, 41(5):501–504, May 1994.

H. Jin and C.X. Ling. Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 17(3):299–310, March 2005.

I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, 1986.

A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. SIAM (Society of Industrial and Applied Mathematics) Press, 1988.

J. B. Keller. Geometrical theory of diffraction. *Journal of the Optical Society of America*, pages 116–130, 1962.

K. T. Kim, D. K. Seo, and H. T. Kim. Radar target identification using one-dimensional scattering centres. In IEE, editor, *Radar, Sonar and Navigation, IEE Proceedings*, volume 148, pages 285–296, October 2001.

R. Klemm. Introduction to space-time adaptive processing. *Electronics and Communication Engineering Journal*, 11:5–12, 1999.

E. F. Knott, J. F. Shaeffer, and M.T. Tuley. *Radar Cross Section*. Artech House, 1985.

Air Force Research Laboratory. www.sdms.afrl.af.mil/datasets/mstar.

V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. *English translation in Soviet Physics Doklady*, 10(8):707–710, 1966.

Hsueh-Jyh Li, Yung-Deh Wang, and Long-Huai Wang. Matching score properties between range profiles of high-resolution radar targets. *IEEE Transactions on Antennas and Propagation*, 44(4):444–452, April 1996.

Xuejun Liao, P. Runkle, Yan Jiao, and L. Carin. Identification of ground targets from sequential HRR radar signatures. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 2897–2900, May 2001.

Xuejun Liao, Hui Li, and B. Krishnapuram. An M-ary KMP classifier for multi-aspect target classification. In *IEEE International Conference on Acoustics, Speech, and Signal Processing.*, volume 2, pages ii – 61 –4, May 2004.

Guoqing Liu, Shunji Huang, A. Torre, and F. Rubertone. The multilook polarimetric whitening filter (MPWF) for intensity speckle reduction in polarimetric SAR images. *IEEE Transaction on Geoscience and Remote Sensing*, 36:1016–1020, May 1998.

C. G. Looney. *Pattern recognition using neural networks*. Oxford University Press, 1998.

D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, March 2003.

J. Martin and B. Mulgrew. Analysis of the theoretical radar return signal form aircraft propeller blades. In *Proceedings of the IEEE Radar Conference*, 1990.

D. L. Mensa. *High Resolution Radar Cross-Section Imaging*. Artech House, 1991.

R. J. Miller, G. C. Sarno, and D. J. Shephard. Progress in radar recognition of aircraft without using radar-derived databases. In *EMRS DTC 1^{st} Annual Technical Conference*, Edinburgh, May 2004.

A. K. Mishra and B. Mulgrew. Bistatic SAR-ATR: A case study. In *EMRS DTC 3^{rd} Annual Technical Conference*, Edinburgh, July 2006.

J. Misiurewicz, K. Kulpa, and Z. Czekala. Analysis of recorded helicopter echo. In *Radar 97 (Conf. Publ. No. 449)*, pages 449–453, October 1997.

R. A. Mitchell and J. J. Westerkamp. Robust statistical feature based aircraft identification. *IEEE transaction on Aerospace and Electronic Systems*, 35(3):1077–1095, July 1999.

I. Moir and A. Seabridge. *Military Avionics Systems*. John Wiley and Sons, 2006.

R. L. Moses and L. C. Potter. Feature extraction using attributed scattering center models for model-based automatic target recognition. In *DARPA Image Understanding Workshop*, 1997.

MSTAR. Program technology review, November 1996.

D. C. Munson, J.D. O'Brien, and W. K. Jenkins. A tomographic formulation of spotlight-mode synthetic aperture radar. *Proceedings of the IEEE*, 72(8):917–925, August 1983.

S. Musman, D. Kerr, and C. Bachmann. Automatic recognition of isar ship images. *IEEE Transactions on Aerospace and Electronic Systems*, 32(4):1392–1404, October 1996.

W. Nel, J. Tait, R. Lord, and A. Wilkinson. The use of a frequency domain stepped frequency technique to obtain high range resolution on the CSIR X-band SAR system. In *IEEE 6th AFRICON Conference in Africa*, volume 1, pages 327–332, October 2002.

F. Neri. *Introduction to Electronic Defense Systems*. SciTech Publishing, 2^{nd} edition, 2006.

Wu Ning, Wugun Chen, and Xinggan Zhang. Automatic target recognition of ISAR object images based on neural network. In *IEEE NNSP 2003*, volume 1, pages 373–376, December 2003.

L. M. Novak. A comparison of 1-D and 2-D algorithms for radar target classification. In *IEEE International Conference on Systems Engineering*, pages 6–12, 1991.

L. M. Novak. Applications of Synthetic Aperture Radar. In *2005 IEEE International Radar Conference Tutorial*, Arlington (VA), May 2005.

L. M. Novak and M. C. Burl. Optimal speckle reduction in polarimetric SAR imagery. *IEEE Transaction on Aerospace and Electronic Systems*, 26:293–305, 1990.

L. M. Novak and S. R. Hesse. Optimal polarizations for radar detection and recognition of targets in clutter. In *SPIE Proccedisngs on Automatic Object Recognition*, volume 1700, pages 114–118,, 1992.

L. M. Novak, G. R. Benitz, G. J. Owirka, and L. A. Bessette. ATR performance using enhanced resolution SAR. In *SPIE Conf. on Algorithms for Synthetic Aperture Radar Imagery III*, volume 2757, pages 332–337, 1996.

L. M. Novak, S. D. Halversen, G. Owirka, and M. Hiett. Effects of polarization and resolution on SAR ATR. *IEEE Transactions on Aerospace and Electronic Systems*, 33:102–116, 1997.

University of Maryland Batlab Website. Auditory neuroethology laboratory, http://www.bsos.umd.edu/psyc/batlab/index.html.

G. J. Owirka, S. M. Verbout, and L. M. Novak. Template-based SAR ATR performance using different image enhancement techniques. In *SPIE Proc. on Algorithms for Synthetic Aperture Radar Imagery VI*, volume 3721, pages 302–319, August 1999.

R. M. Page. *The Origin of Radar*. Anchor Books, Garden City, NY, 1962.

A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Int. Ed., 1984.

Bingnan Pei and Zheng Bao. Multi-aspect radar target recognition method based on scattering centers and HMMs classifiers. *IEEE Transactions on Aerospace and Electronic Systems*, 41 (3):1067, 1074 2005.

W. Peterson, T. Birdsall, and W. Fox. The theory of signal detectability. *IEEE Transactions on Information Theory,*, 4(4):171–212, September 1954.

E. Piazza. Radar signals analysis and modellization in the presence of JEM application to civilian ATC radars. *IEEE Aerospace and Electronic Systems Magazine*, 14:35–40, 1999.

M. Plonus, R. Williams, and S. Wang. Radar cross section of curved plates using geometrical and physical diffraction techniques. *IEEE Transactions on Antennas and Propagation*, 26 (3):488–493, May 1978.

L. C. Potter, Da-Ming Chiang, R. Carriere, and M. J. A. Gerry. A GTD-based parametric model for radar scattering. *IEEE Transactions on Antennas and Propagation*, 43(10):1058–1067, October 1995.

D. W. Purnell, E. C. Botha, and C. Nieuwoudt. Classification of two ship targets using radar

backscatter. In *Proceedings of the 1994 IEEE South African Symposium on Communications and Signal Processing*, pages 138–145, September 1998.

L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, February 1989.

A. W. Rihaczek. Radar resolution of ideal point scatterers. *IEEE Transaction on Aerospace and Electronic Systems*, 32:842–845, 1996.

A. W. Rihaczek and S. J. Herschowitz. *Theory and Practice of Radar Target Identification.* Artech House, 2000.

R. Ross. Radar cross section of rectangular flat plates as a function of aspect angle. *IEEE Transactions on Antennas and Propagation*, 14(3):329–335, May 1966.

P. Runkle, P. K. Bharadwaj, L. Couchman, and L. Carin. Hidden markov models for multiaspect target classification. *IEEE Transactions on Signal Processing*, 47(7):2035–2040, July 1999a.

P. R. Runkle, P. K. Bharadwaj, L. Couchman, and L. Carin. Hidden Markov models for multiaspect target classification. *IEEE Transactions on Signal Processing*, 47(7):2035–2040, July 1999b.

F. Sadjadi. Improved target classification using optimum polarimetric SAR signatures. *IEEE Transactions on Aerospace and Electronic Systems*, 38:38–49, 2002.

J. Salazar, M. Robinson, and M. R. Azimi-Sadjadi. Multi-aspect discrimination of underwater mine-like object objects using hidden markov models. In *Proceedings of the Oceans '02 MTS/IEEE*, volume 1, pages 46–53, October 2002.

O. S. Sands and F. D. Garber. Syntactic classification of radar measurements of commercial aircraft. In *Proceedings of the 1989 IEEE National Radar Conference*, pages 158–163, March 1989.

R. A. Saville-Sneath. *Aircraft Recognition.* Penguin Books, 2006, (First Published 1941).

R. 0. Schmidt. Multiple emitter location and signal parameter estimation. In *RADCSpectral Estimation Workshop*, pages 243–258, Rome, NY, 1979.

R. Schumacher and J. Schiller. Non-cooperative target identification of battlefield targets - classification results based on SAR images. In *2005 IEEE International Radar Conference*, pages 167–172, Arlington (VA), May 2005.

Ji Shihao, Liao Xuejun, and L. Carin. Adaptive multiaspect target classification and detection with hidden markov models. *IEEE Sensors Journal*, 5(5):1035–1042, October 2005.

G. A. Showman, M. A. Richards, and K. J. Sangston. Comparison of two algorithms for correcting zero-Doppler clutter in turntable isar imagery. In *Proceedings of the Conference on Signals, Systems & Computers*, volume 1, pages 411–415, 1998.

J. A. Simmons, P.A. Saillant, and S. P. Dear. Through a bat's ear. In *IEEE Spectrum*, volume 29, pages 46–48, March 1992.

M. I. Skolnik. *Radar Handbook (2nd Edition)*. McGraw-Hill, 1990.

H. C. Stankwitz and M. R. Kosek. Super-resolution for SAWISAR RCS measurement using spatially variant apodization (super-SVA). In *Proc. Antenna Measurement and Techniques Association (AMTA) Symp.*, November 1995.

H. C. Stankwitz, R. J. Dallaire, and J. R. Fienup. Nonlinear apodization for sidelobe control in SAR imagery. *IEEE Transaction on Aerospace and Electronic Systems*, 31:267–279, 1996.

G. W. Stimson. *Introduction to Airborne Radar*. SciTech Publishing, 2 edition, 1998.

R. J. Sullivan. *Radar Foundations for Imaging and Advanced Concepts*. SciTech Publishing, 2004.

S. S. Swords. *Technical History of the Beginnings of Radar*. IET, June 1986.

P. Tait. Introduction to radar target recognition. *IEE, Publishing*, 2006.

S. Theodoridis and K. Koutroumbas. *Pattern Recognition (3rd edition)*. Elsevier, 2006.

J. Tsao and B. D. Steinberg. Reduction of sidelobe and speckle artifacts in microwave imaging: the CLEAN technique. *IEEE Transaction on Antennas and Propagation*, 36:543–556, 1988.

D. Vernon. *Machine Vision - Automated Visual Inspection and Robot Vision*. Prentice-Hall, 1991.

M. Vespe. Multi-perspective radar target classification. University College London, April 2005. MPhil/PhDTransfer Thesis.

M. Vespe, C. J. Baker, and H. D. Griffiths. Multi-perspective radar target classification. In *EMRS DTC 1st Annual Technical Conference*, Edinburgh, May 2004.

M. Vespe, C. J. Baker, and H. D. Griffiths. Radar target classification using multiple perspectives. In *EMRS DTC 2nd Annual Technical Conference*, Edinburgh, June 2005a.

M. Vespe, C. J. Baker, and H. D. Griffiths. Multi-perspective target classification. In *2005 Int. IEEE Radar Conference*, pages 877–882, Arlington (VA), May 2005b.

M. Vespe, C. J. Baker, H. D. Griffiths, P. Lombardo, and D. Pastina. Feature extraction for SAR target classification. In *London Communications Symposium*, London, 2005c.

M. Vespe, C. J. Baker, and H. D. Griffiths. Outline structural representation for radar target classification based on non-radar templates. In *CIE Int. Radar Conference*, October 2006a.

M. Vespe, C. J. Baker, and H. D. Griffiths. Perspective aperture dependency of netted radar target classification. In *EMRS DTC 3rd Annual Technical Conference*, Edinburgh, July 2006b.

M. Vespe, C. J. Baker, and H. D. Griffiths. Frequency diversity vs large bandwidth reconstruction: information content for netted sensor ATR using ISAR images. In *International Conference on SAS and SAR*, Lerici (Italy), September 2006c.

M. Vespe, C. J. Baker, and H. D. Griffiths. Multi perspective target classification. In *SET-095 NATO Specialist Meeting on Bistatic and Multistatic Radar and Sonar Systems*, La Spezia, July 2006d.

M. Vespe, C. J. Baker, and H. D. Griffiths. Aspect dependent drivers for multi-perspective target classification. In *2006 IEEE Radar Conference*, pages 256–260, Verona (NY), April 2006e.

M. Vespe, C. J. Baker, and H. D. Griffiths. Node location for netted radar target classification. In *2006 International Waveform Diversity & Design Conference*, pages 256–260, Hawaii, January 2006f.

M. Vespe, C. J. Baker, H. D. Griffiths, P. Lombardo, and D. Pastina. Impact of angular processing techniques and cross-range resolution on 2-D radar target classification. In *6th European Conference on Synthetic Aperture Radar*, Dresden, May 2006g.

M. Vespe, G. Jones, and C. J. Baker. Echolocating bats for autonomous navigation. *Journal of the Acoustical Society of America*, Paper in revision.

M. Vespe, C. J. Baker, and H. D. Griffiths. Multi-signature radar target classification. *Special Issue of IET: Radar, Sonar and Navigation*, in revision.

M. D. De Vore and J. A. O'Sullivan. Performance complexity study of several approaches to automatic target recognition from SAR images. *IEEE Transactions on Aerospace and Electronic Systems*, 38(2):632–648, April 2002.

D. R. Wehner. *High resolution radar*. Artech House, 1995.

A. J. Wilkinson, R. T. Lord, and M. R. Inggs. Stepped-frequency processing by reconstruction of target reflectivity spectrum. In *Communications and Signal Processing, '98*, pages 101–104, 1998.

R. Williams, J. Westerkamp, D. Gross, and A. Palomino. Automatic target recognition of time critical moving targets using 1d high range resolution (HRR) radar. In *IEEE Aerospace and Electronic Systems Magazine*, volume 15, pages 37–43, April 2000.

K. Woods and K. W. Bowyer. Generating ROC curves for artificial neural networks. *IEEE Transactions on Medical Imaging*, 16(3):329–337, June 1997.

R. Wu, Q. Gao, J. Liu, and H. Gu. Atr scheme based on 1-D HRR profiles. In *Electronics Letters*, volume 38, pages 1586–1588. IEEE, November 2002.

Liao Xuejun, P. Runkle, and L. Carin. Identification of ground targets from sequential high-range-resolution radar signatures. *IEEE Transactions on Aerospace and Electronic Systems*, 38(4):1230–1242, October 2002.

A. Zyweck and R.E. Bogner. Radar target classification of commercial aircraft. *IEEE Transactions on Aerospace and Electronic Systems*, 32(2):598–606, April 1996.

A. Zyweck and R.E. Bogner. Coherent averaging of range profiles. In *Record of the IEEE 1995 International Radar Conference*, pages 456–461, May 1995.